

**ANALYSIS OF ERROR-RELATED POTENTIALS IN P300 AND  
MOTOR IMAGERY BASED BRAIN COMPUTER INTERFACES**

by  
**Abdullahi Adamu**

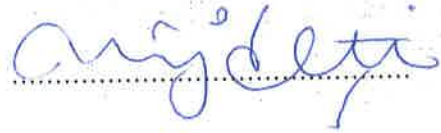
Submitted to the Graduate School of Engineering and Natural Sciences  
in partial fulfillment of  
the requirements for the degree of  
Master of Science

**Sabancı University**  
**August 2016**

ANALYSIS OF ERROR-RELATED POTENTIALS IN P300 AND MOTOR  
IMAGERY BASED BRAIN COMPUTER INTERFACES

APPROVED BY

Assoc. Prof. Dr. Müjdat Çetin  
(Thesis Supervisor)



Assoc. Prof. Dr. Devrim Ünay



Assist. Prof. Dr. Murat Kaya Yapıcı



DATE OF APPROVAL: 25/08/2016

© Abdullahi Adamu 2016

All Rights Reserved

*to my loved ones...*



## Acknowledgments

This thesis is the product of a combined effort originating from both within and outside my professional life. The most dominant figure that has influenced me during the past 2 years is my academic supervisor and mentor, Müjdat Çetin. He has been there to ensure that I have made the most out of good times and especially, the difficult times. Working under his guidance has taught me many things such as professionalism; the ability to work with different kinds of people and set high standards in all that I do, kindness; the ability to look past the faults in others and instead provide them with support and encouragement, humbleness; the realization that in the end, no one is above all and that we are all in this together, doing the best that we can to leave a better world for the next generation.

When I started conducting research, Sumeyra Ummuhan Demir Kanik has been with me every step of the way. She has given me a better understanding of how research should be done, what questions should be asked, and how we can make sense of the results we obtain. She has also been instrumental when I began writing my thesis. Her invaluable suggestions, corrections, and patience with me has enabled me to write a thesis I am very proud of, and I hope that she is too.

Working in the SPIS and VPA labs has been a privilege and part of that is thanks to Osman Rahmi Ficici who works tirelessly behind the scenes to ensure that we not only have what we need, but also that we also have the best available.

Among five graduates this year from our BCI group, I joined the latest. Despite this, I have been fortunate to easily get up to speed thanks to the unending support of my friends from the BCI group. Ozan Özdenizci, Sezen Yağmur Gunay, Mastaneh Torkamani Azar, and Majed Elwardy have provided a working environment that has

always been rich with ideas and assistance in every form. I have also had the privilege of benefiting from the experience of colleagues outside the BCI group. Burak Alver, Muhammad Usman Ghani, Oğuzcan Zengin, and Naeimeh Atabakilachini have broadened my mind with new perspectives that have helped shape the way I think over the past 2 years.

All work and no play makes Jack a dull boy. This is why in addition to my aforementioned colleagues, I will like to express my gratitude to the friends that have given me so much to be grateful to; Abba Ibrahim Ramadan, a friend of six years with whom I share so many memories, hopes, and dreams; Nassur Mohammed Ramadhan, a fellow African who always reminds me not to take myself too seriously with his refreshing character; Wisdom Chukwunwike Agboh, a fellow Nigerian who always reminds me of when to take myself seriously.

To friends who have been by my side even from the farthest of distances, I am most grateful for the impact you have had upon me. Shettima Sani Dambatta, Abubakar Isa Adamu, Oyewole Efunbajo, John Omole, Adamu Abdullahi, and many others who have supported me have my deepest gratitude.

I am the person I am today because of the nurture and support of my family. They have gone beyond their ways to ensure that I have good prospects for the future while remembering what is truly important – who I am and where I come from. My gratitude to them is immeasurable.

I am especially grateful to TÜBİTAK for their financial support throughout my graduate study and to Sabancı University giving me the opportunity to live in an environment that has nurtured me both academically and as a person.

# ANALYSIS OF ERROR-RELATED POTENTIALS IN P300 AND MOTOR IMAGERY BASED BRAIN COMPUTER INTERFACES

Abdullahi Adamu

EE, M.Sc. Thesis, 2016

Thesis Supervisor: Müjdat Çetin

**Keywords:** electroencephalogram, brain-computer interfaces, error related potentials, adaptation

## Abstract

Brain Computer Interface (BCI) systems aim to generate alternative communication pathways for people with disabilities by extracting information directly from the brain. Increasing interest in this field of study has enabled patients to use electroencephalography (EEG) in controlling word processing software such as the P300 speller and prostheses using motor imagery through EEG. Despite achieving successful real-time implementations in these applications, Brain Computer interfaces are subject to errors when interpreting the user's intent. One way of reducing this is by using the Error Related Potential (ErrP). These are signals generated by a person when an error occurs in a BCI system. The knowledge that an error has occurred in a BCI could perhaps be used in strengthening the decision making process of the BCI. Our work aims to understand the effect of different types of user involvement has on ErrP waveforms and classification performance in P300 and motor imagery based BCI experiments. Particularly, we collect data in three different settings for both P300 and motor imagery based BCIs and provide an analysis of this data using signal processing and machine learning techniques. We also show how results obtained from the motor imagery based experiments can be used as a basis for a BCI system where motor imagery and Error

Related Potentials are classified simultaneously. Furthermore, preliminary experiments have been done to classify motor imagery and ErrP in this joint motor imagery and ErrP detection system. We have also investigated the effect of changes in trial frequency on ErrP classification performance in motor imagery based BCI systems.

# P300 VE HAYALİ MOTOR HAREKETİNE DAYALI BEYİN BİLGİSAYAR ARAYÜZLERİNDE HATAYA DAYALI POTANSİYELLERİN ANALİZİ

Abdullahi Adamu

EE, Yüksek Lisans Tezi, 2016

Tez Danışmanı: Müjdat Çetin

**Anahtar Kelimeler:** elektroensefalografi, beyin bilgisayar arayüzü, hataya dayalı potansiyel, uyarılama

## Özet

Beyin bilgisayar arayüzü (BBA) sistemleri, beyinden belli sinyalleri toplayarak engelliler için alternatif bir haberleşme yöntemi sağlamayı hedefler. Bu alana ilginin artmasıyla beraber hastaların elektroensefalografi (EEG) sinyalleriyle P300 heceleyici gibi kelime işleyen ve hayali motor hareketleri kullanarak protez kontrol eden sistemler kullanmaları mümkün olmuştur. Bu uygulamaların gerçek zamanlı kullanmasında başarılar elde edilse de, beyin bilgisayar arayüzlerinin kullanıcının niyetini yorumlaması hatalı olabilmektedir. Bu hataları azaltmanın bir yolu hataya ilişkin potansiyelleri (ErrP) kullanmaktır. ErrP, BBA sistemlerinde bir hata meydana geldiğinde beyinde üretilen sinyaldir. BBA’da bir hata oluştuğunun bilgisi BBA’nın karar alma mekanizmasını güçlendirmekte kullanılabilir. Bizim çalışmamızın amacı, farklı kullanıcı katılım düzeneklerinin ErrP dalgalarına ve BBA deneylerinde sınıflandırma performansına olan etkilerini anlamaktır. Bu amaçla, P300 ve hayali motor tabanlı BBA’lar için üç farklı düzenekte veri topladık, ve sinyal işleme ve makine öğrenme teknikleri kullanarak bu verileri analiz ettik. Ayrıca, hayali motor deneylerinden elde edilen sonuçların hayali motor ve ErrP sinyallerinin aynı anda sınıflandırıldığı BBA sistemi için bir temel olarak kullanabileceğini gösterdik. Bunun yanında, hayali motor ve ErrP sinyallerinin eş za-

manlı kaydedilerek sınıflandırmaları amacıyla ön deneyler yaptık. Son olarak deneme frekansındaki değişimin, hayali motor tabanlı BBA sistemlerindeki ErrP sınıflandırma performansı üzerine etkisini inceledik.

# Table of Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Özet</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Scope and Motivation . . . . .	3
1.2 Contributions . . . . .	4
1.3 Outline . . . . .	5
<b>2 Background</b>	<b>7</b>
2.1 BCI System . . . . .	8
2.2 EEG Signals . . . . .	10
2.2.1 Sensorimotor Rhythms . . . . .	13
2.2.2 Event Related Potential . . . . .	14
2.2.3 Error Related Potentials . . . . .	16
2.3 Classification Methods . . . . .	19
2.3.1 Linear Discriminant Analysis . . . . .	19
2.3.2 Support Vector Machines . . . . .	20
2.3.3 Gaussian Mixture Models . . . . .	20
2.4 Adaptation in Brain Computer Interfaces . . . . .	22
2.4.1 Nonstationarity in EEG . . . . .	22
2.4.2 P300 Based BCI . . . . .	23
2.4.3 Motor Imagery Based BCI . . . . .	25
<b>3 Analysis of ErrP in P300 based Brain Computer Interfaces</b>	<b>28</b>
3.1 Experimental Description . . . . .	29
3.1.1 Data Processing . . . . .	31
3.1.2 Classification . . . . .	32
3.2 P300 Observe . . . . .	32
3.2.1 Experimental Description . . . . .	32
3.2.2 Results . . . . .	33
3.3 P300 Control . . . . .	34
3.3.1 Experimental Description . . . . .	34

3.3.2	Results . . . . .	34
3.4	P300 EEG . . . . .	35
3.4.1	Results . . . . .	36
3.5	Discussion . . . . .	37
3.5.1	Waveform analysis . . . . .	37
3.5.2	Performance analysis . . . . .	40
3.6	Summary . . . . .	45
<b>4</b>	<b>Analysis of ErrP in Motor Imagery based Brain Computer Interfaces</b>	<b>46</b>
4.1	Processing Preexisting BCI Datasets . . . . .	47
4.1.1	Experimental Description . . . . .	47
4.1.2	Results . . . . .	47
4.2	Designed Interface for MI ErrP . . . . .	48
4.2.1	Data Processing . . . . .	51
4.2.2	Classification . . . . .	51
4.3	Motor Imagery Observe . . . . .	51
4.3.1	Experimental Description . . . . .	51
4.3.2	Results . . . . .	51
4.4	Motor Imagery Control . . . . .	53
4.4.1	Experimental Description . . . . .	53
4.4.2	Results . . . . .	53
4.5	Motor Imagery EEG . . . . .	54
4.5.1	Experimental Description . . . . .	54
4.5.2	Results . . . . .	54
4.6	Discussion . . . . .	56
4.6.1	Waveform analysis . . . . .	56
4.6.2	Performance analysis . . . . .	58
4.7	ErrP across P300 and motor imagery based BCI experiments . . . . .	63
4.8	Summary . . . . .	66
<b>5</b>	<b>Analysis and Design of a Joint Motor Imagery and ErrP-detection System</b>	<b>68</b>
5.1	Detection of error related potentials . . . . .	69
5.1.1	Design of Error Detection System . . . . .	69
5.1.2	Results . . . . .	72
5.2	Motor Imagery Classification . . . . .	77
5.2.1	Interface Design and Data Processing . . . . .	77
5.2.2	Results . . . . .	78
5.3	Joint Motor Imagery and ErrP Detection . . . . .	79
5.3.1	Interface Design and Data Processing . . . . .	79
5.4	Results . . . . .	80
5.5	Summary . . . . .	82



<b>6</b>	<b>Conclusion and Future Work</b>	<b>84</b>
6.1	Conclusion . . . . .	84
6.2	Future Work . . . . .	87
	<b>Bibliography</b>	<b>88</b>

## List of Figures

1.1	An illustration of a typical BCI system . . . . .	2
1.2	An illustration of a BCI system integrated that uses error related potentials. . . . .	4
2.1	A basic EEG based BCI setup. A user sits in front of a computer screen and performs a series of tasks. During this period, EEG signals are recorded from the scalp and fed into the EEG amplifier. Amplified EEG signals are recorded and stored using special BCI software. During an experiment, feedback is usually presented to the user either visually or acoustically. . . . .	9
2.2	Real time raw EEG signals recorded by the ActiView software. . . . .	10
2.3	The electrodes used to record surface electrical activity of the brain. . . . .	11
2.4	Gel is first put on the surface of a subject's head before electrodes are placed. It is used to ensure the conductivity between the scalp and the electrodes. . . . .	12
2.5	The standard 10-20 system showing electrode placement locations. . . . .	13
2.6	The overall average waveform of the error minus correct responses obtained from the work of Chavarriaga et al. [1]. . . . .	17
3.1	The main page of the P300 related experiments where experiment-specific settings are entered. . . . .	29
3.2	Target Phase of P300 based BCI experiments. Subjects are shown a random letter for one second. . . . .	30

3.3	Selection Phase of P300 based BCI experiments. Rows and columns randomly flash during this phase. . . . .	30
3.4	Feedback Phase of P300 based BCI experiments. A feedback is presented to the subjects and they are instructed to notice if it is the same as the letter presented during the target phase. . . . .	31
3.5	Average waveforms over all subjects for correct (blue), error (red) and error-minus-correct (yellow) samples for all three protocols. First column represents results obtained from the FCz electrode and the second column represents results obtained from the Cz electrode. Top row represents results for the <i>Observe</i> protocol, while the middle row represents the <i>Control</i> protocol and the bottom row represents the <i>EEG</i> protocol. . . . .	38
3.6	This figure shows the mean error minus correct waveforms for all three protocols as recorded from the FCz electrode. Blue represents <i>EEG</i> protocol, red represents <i>Observe</i> protocol and yellow represents <i>Control</i> protocol. . . . .	39
3.7	This figure presents mean accuracies obtained over all subjects and in all three protocols for four different classifiers; SVM, LDA, Decision Tree and GMM. . . . .	41
4.1	The Start Menu of the Motor imagery related experiments. . . . .	49
4.2	The Resting Phase of the Motor Imagery related experiments. . . . .	49
4.3	The Stimulus Phase of the Motor Imagery related experiments. . . . .	50
4.4	One trial in the motor imagery based BCI experiments. . . . .	50
4.5	This figure presents overall average waveforms computed for correct (blue), error (red), and error-minus-correct (yellow) samples for two different electrodes and for all three protocols. From top to bottom, the rows are arranged in the following order: <i>Observe</i> , <i>Control</i> , <i>EEG</i> . The left column represents data from the FCz electrode and the right column represents data from the Cz electrode. . . . .	56

4.6	This figure shows the mean error minus correct waveforms for all three protocols recorded from the FCz electrode. Blue waveforms represent the <i>EEG</i> protocol, red represents <i>Observe</i> protocol and yello represents the <i>Control</i> protocol. . . . .	57
4.7	Average ErrP classification results over all subjects for all three protocols. Classifiers used include SVM, LDA, Decision Tree and GMM. . . . .	59
4.8	Average error minus correct ErrP waveforms for P300 and motor imagery based BCI experiments for all three protocols. Blue lines show the difference waveform for motor imagery experiments scaled down by a factor of 4. Red lines show the difference waveform for P300 based BCI experiments without scaling. . . . .	64
4.9	Difference waveforms of all three runs in motor imagery based experiments for all protocols. Blue lines represent difference waveforms obtained in the first runs, red lines correspond to waveforms obtained in the second runs, and yellow lines correspond to waveforms obtained in the third runs. . . . .	65
4.10	Difference waveforms of all three runs in P300 based experiments for all protocols. Blue lines represent difference waveforms obtained in the first runs, red lines correspond to waveforms obtained in the second runs, and yellow lines correspond to waveforms obtained in the third runs. . . . .	66
5.1	Interface of the one step protocol. . . . .	70
5.2	Interface of the three step protocol. . . . .	71
5.3	Interface of the six step protocol. . . . .	72
5.4	Average ErrP classification performance over all subjects for all three protocols. Classifiers used include SVM, LDA, Decision Tree and GMM classifiers. . . . .	74
5.5	One trial in the preliminary motor imagery experiments. . . . .	77

5.6	Electrodes used in the joint motor imagery and ErrP detection protocol. Blue colors represent electrodes used in classification and yellow represents electrodes used for referencing. . . . .	78
5.7	Error minus correct waveform averaged over all subjects for the FCz electrode. . . . .	81
5.8	ErrP classification performance for LDA, SVM, Decision Tree, and GMM classifiers averaged over all subjects for motor imagery <i>EEG</i> (blue) and the joint motor imagery and ErrP detection (red) protocols. . . . .	82

## List of Tables

3.1	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the <i>Observe</i> protocol in P300 based BCI experiments. For each electrode, the accuracy of classifying correct and error samples are presented. . . . .	34
3.2	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the <i>Control</i> protocol in P300 based BCI experiments. For each electrode, the accuracy of classifying correct and error samples are presented. . . . .	35
3.3	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the <i>EEG</i> protocol in P300 based BCI experiments. For each electrode, the accuracy of classifying correct and error samples are presented. . . . .	37
3.4	Maximum positive and negative deflections for the average error-minus-correct waveforms obtained for all three protocols. . . . .	40
3.5	MANOVA test results for pairwise combinations of the three protocols in the P300 based experiments. . . . .	42
3.6	Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the <i>Observe</i> protocol of P300 based BCIs. . . . .	43
3.7	Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the <i>Control</i> protocol of P300 based BCIs. . . . .	44

3.8	Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the <i>EEG</i> protocol of P300 based BCIs. . . . .	45
4.1	ErrP classification results obtained by Millan et. al. [1] and the results we have obtained using our codes on the same data. . . . .	48
4.2	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the <i>Observe</i> protocol. For each electrode, the accuracy of classifying correct and error samples are presented. . .	52
4.3	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the <i>Control</i> protocol. For each electrode, the accuracy of classifying correct and error samples are presented. . .	54
4.4	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the <i>EEG</i> protocol. For each electrode, the accuracy of classifying correct and error samples are presented. . . . .	55
4.5	Maximum positive and negative deflections computed from the waveforms provided in Figure 4.6. . . . .	58
4.6	MANOVA test results for pairwise combinations of the three protocols in the motor imagery based experiments. . . . .	60
4.7	Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the <i>Observe</i> protocol of motor imagery based BCIs. . . . .	61
4.8	Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the <i>Control</i> protocol of motor imagery based BCIs. . . . .	62
4.9	Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the <i>EEG</i> protocol of motor imagery based BCIs. . . . .	63
5.1	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the one step protocol. . . . .	73

5.2	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the three step protocol. . . . .	73
5.3	ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the six step protocol. . . . .	74
5.4	MANOVA test results for testing the significance of the difference between the accuracies obtained for the one step, three step, and six step protocols. This test is performed for four different classifiers; SVM, LDA, GMM, and Decision Tree. . . . .	75
5.5	Optimal electrodes and their performances for all subjects and classifiers in the one step protocol. . . . .	75
5.6	Optimal electrodes and their performances for all subjects and classifiers in the three step protocol. . . . .	76
5.7	Optimal electrodes and their performances for all subjects and classifiers in the six step protocol. . . . .	76
5.8	Subject performance in motor imagery experiments. . . . .	79
5.9	Motor imagery performance in the joint motor imagery and ErrP detection experiments. . . . .	80
5.10	ErrP classification results in the joint motor imagery and ErrP detection experiments. . . . .	80



## Chapter 1

# Introduction

The prospect of using electrical activity recorded from the brain as an alternative means of communication for humans has been gaining research interest over the past decades [2]. One of the main motivating factors that drives this effort is an increasing desire to help people that have lost the ability to communicate effectively, such as stroke patients and patients suffering from amyotrophic lateral sclerosis (ALS). Nearly 15 million people worldwide suffer from stroke [3] and over the past 50 years, 1 to 7 of every 100,000 adults worldwide is estimated to have ALS at any given time [4].

The physically disabled currently benefit from a wide range of solutions. Physical rehabilitation for example, is a very common way of helping patients restore vital movements in their body. Physical rehabilitation can be challenging on the patient and sometimes, it might not produce desired results. This has researchers to question the effectiveness of physical rehabilitation on disabled patients. One study [5] found no or insufficient evidence on the basis of functional outcome for various physical rehabilitation protocols. In a way, Brain Computer Interfaces (BCI) become relevant because these systems aim to extract useful information directly from human brain activity and then use this information to ease communication and rehabilitation for disabled patients.

We can define a now Brain Computer Interface as a system that provides alternate communication and control channels for the human brain that are independent of regular channels such as peripheral nerves and muscles [6]. In that sense, when a patient has

difficulty controlling their muscles for any reason, the BCI can be used to bypass the neural connection and directly infer what the person is trying to do. This can be done by recording brain activities and interpreting these activities using state-of-the-art technology. Figure 1.1 shows a typical BCI system. Electrical activity from the brain is recorded from a user using special electronic devices. These signals are preprocessed using various signal processing techniques after which task relevant features can be extracted and classified using machine learning techniques. The classification result can be presented as an output to the user via a feedback mechanism, typically visually or auditory.

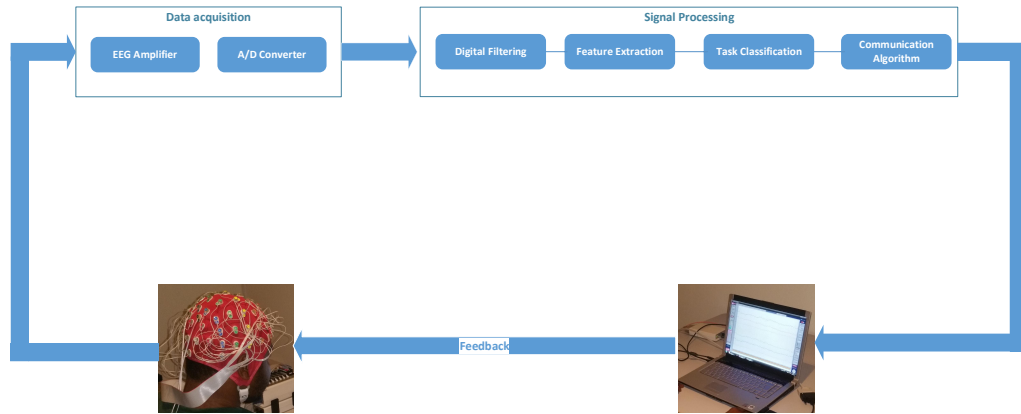


Figure 1.1: An illustration of a typical BCI system

Recording high quality brain signals is important in BCI research. Given the complex nature of the human brain involving millions of neural activity, it is important to extract brain signals with a signal-to-noise ratio that is as high as possible. The best way of doing this is by using invasive methods. Invasive methods are methods of recording brain activity from electrodes implanted beneath the skull through a surgical operation. These methods produce signals with high signal-to-noise ratios due to their proximity to the surface of the brain. However, these methods are inconvenient because of the surgery that is required to plant these electrodes. Other methods that avoid this problem exist and are known as non-invasive methods. In contrast to invasive methods, non-invasive methods record brain activity with sensors located outside the scalp. These methods are more convenient but come with at a cost of reduced signal-to-noise

ratio.

A popular way of recording brain signals is electroencephalography (EEG). EEG is so popular not only because it is non-invasive and cheap, several studies have also shown that there is a link between recorded EEG and mental tasks [7, 8]. The existence of such links has opened up numerous research opportunities in a collaborative effort between fields such as engineering and neuroscience.

## 1.1 Scope and Motivation

The central focus of this thesis is a special type of EEG potential known as the error related potential (ErrP). Knowledge of EEG helps to provide a good perspective on why ErrPs are important. In BCIs, EEG signals are usually preprocessed to suit computational requirements. Following this, certain features of the data are extracted to make decisions. These features vary for different experiments. In relation to this case, it has been found that subjects produce a certain potential in response to an error when using a BCI system. This error can be an error made by the BCI by incorrectly interpreting the subject's intent, or it could even come from a realization from the subject that they had in fact committed an error. The presence of error related potentials can be useful in updating classifiers so similar occurrences can be prevented in the future. Figure 1.2 shows a typical BCI that uses error related potentials. In this setting, features are extracted for the BCI task and ErrP classification. The output of the ErrP classifier and task classifier are used to continuously update the decision making process of the BCI.

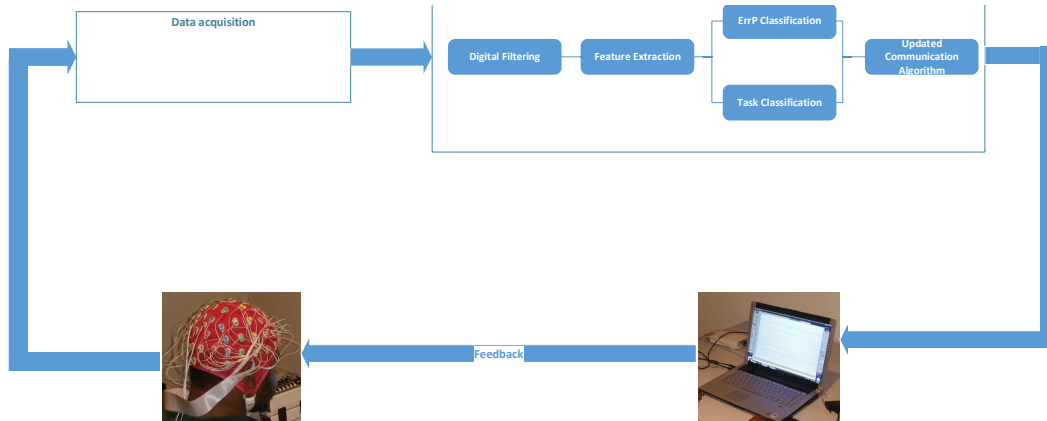


Figure 1.2: An illustration of a BCI system integrated that uses error related potentials.

Based on the nature of error related potentials, a new direction for the BCI group in our institution is finding ways to improve the current BCI systems – motor imagery and P300 based BCIs – by using these potentials. Before that, we wanted to understand the nature of these potentials in different contexts and analyze the performance of ErrP classification for these contexts.

Once that is achieved, the second step will be the integration of error related potentials into the systems and devising strategies upon which they can be used to improve performance.

## 1.2 Contributions

We propose an analysis of error related potentials in P300 and motor imagery based BCI protocols with the aim of eventually using the knowledge gained to improve performance in these protocols. Our analysis is motivated by answering the following questions.

1. How do different contexts affect error related potentials in P300 based BCIs?
2. How do different contexts affect error related potentials in motor imagery based BCIs?

3. Can we classify motor imagery and error related potentials in a single joint experiment?

In this thesis, we design three different protocols in both P300 and motor imagery based Brain Computer Interfaces. These protocols represent three different contexts: *Observe* – where a subject does not have control on the BCI, *Control* – where a subject presses keys on the keyboard to control the BCI, and *EEG* – where a subject uses their brain signals to control the BCI. We compute the waveforms generated by these protocols for both P300 and motor imagery based BCIs and provide analyses of these waveforms. Another contribution we provide is a performance analysis of ErrP in all protocols for four different classifiers; Support Vector Machines, Linear Discriminant Analysis, Decision Tree, and Gaussian Mixture Model. Additionally, we analyze how error related potentials differ across P300 and motor imagery based experiments.

One more contribution of this thesis is the analysis of the effect of changes in frequency of trials on classification performance of four different classifiers, namely Support Vector Machines, Linear Discriminant Analysis, Decision Tree, and Gaussian Mixture Model. This analysis provides information on the robustness of these classifiers to changes in EEG.

The final contribution of this thesis is the design, implementation, and analysis of a system that classifies motor imagery and error related potentials in a single experiment.

## 1.3 Outline

The thesis is organized as follows.

In Chapter 2, we provide background information on BCI systems, EEG signals such as sensorimotor rhythms and error related potentials, classification methods and adaptation in Brain Computer Interfaces.

In Chapter 3 we describe the three protocols we have designed for ErrP analysis in P300 based BCIs. We provide feature extraction and classification techniques used in this work. We end the chapter by presenting the results we have obtained.

Chapter 4 presents three protocols designed to analyze ErrP in motor imagery based BCIs. Feature extraction and classification techniques are provided and the results related to this work follows.

In Chapter 5, we first describe two preliminary studies we have done. The first is a study of three different motor imagery protocols designed to analyze how ErrP performance is affected by changes in the frequency of trials. The second is a study on two subjects to determine motor imagery performance on our designed interface. Results of these preliminary studies are also presented in this chapter. A description of the joint motor imagery and ErrP detection system is then presented and the chapter ends by providing the results obtained in this work.

Chapter 6 provides a summary of the results obtained in Chapters 3, 4 and 5. It also includes suggestions for possible research directions that can be taken based on the work that has been done in this thesis.

## Chapter 2

# Background

This chapter provides background information and literature review on concepts such as BCI systems, EEG signals, and signal processing techniques used in this area.

One of the main forces that drive research on Brain Computer Interfaces is the prospect of finding solutions that improve the lives of people with different neurological disorders [9, 10, 11]. It offers a way to help stroke patients regain control of their limbs [12, 13]. It seeks a way to help patients suffering from ALS send an email without the assistance of another person [14]. It also offers a way to help people that have lost their ability to speak to still communicate with the outside world [15].

To accomplish this, understanding the fundamental aspects that govern brain activity is important. This is why numerous research efforts have been made to have a better understanding of the human brain. By doing so, scientists had hoped to discover components of human brain activity that can be directly translated into certain actions by the user. Mapping signal components to their corresponding actions enables the BCI to determine what a person is trying to do just by observing their brain activity.

Brain signal acquisition can be classified into two different categories; invasive and noninvasive [2]. Invasive methods involve placing sensors under the skull. This method produces signals of very high quality but the cost of doing so is high for two reasons. The first reason is that a brain surgery is needed to implant these sensors while the second reason is that these sensors can only last a limited amount of time [16]. Non-invasive methods on the other hand measure brain activity from the surface of the scalp.

This method is very cheap relative to the invasive methods but the trade off is that the quality of the signal is much lower compared to the invasive methods. Examples of non-invasive methods include magnetoencephalography (MEG) [17], positron emission tomography (PET) [18] and functional magnetic resonance imaging (fMRI) [19], and electroencephalography (EEG) [20].

BCI research has already shown that this form of communication is possible. In 1973, Vidal. et al. successfully used Visual Evoked Potentials (VEP) recorded over the visual cortex to infer the direction of a user's gaze [21]. Seven years later, Birbaumer et al. showed that it was possible for users to modulate their Slow Cortical Potentials to move cursors on a computer screen [22].

## 2.1 BCI System

Figure 2.1 gives a general idea an EEG based BCI setup. EEG signals recorded from the scalp are first processed by the EEG Amplifier. This amplifies the EEG signal and also to gets rid of DC and unnecessary high frequency components from the signal. The next step is converting the signal from analog form to a digital form that can be processed by a computing device, typically a computer.

Figure 2.2 shows raw EEG signals recorded from FC1, FC2, C1, C2, Cz, Fz, and CPz electrodes. It should be noted that most of the time, the signals are preprocessed to get rid of certain frequency components before features can be extracted. The nature of the signal processing techniques performed in this case depends on the type of BCI application used.

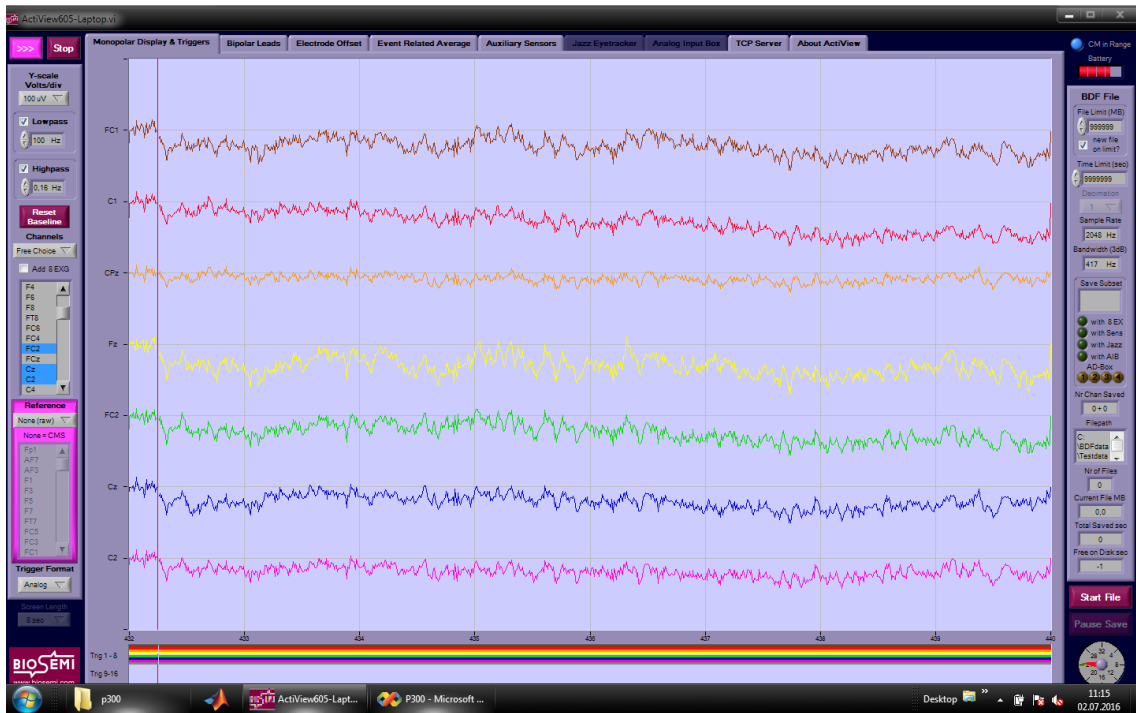




Figure 2.1: A basic EEG based BCI setup. A user sits in front of a computer screen and performs a series of tasks. During this period, EEG signals are recorded from the scalp and fed into the EEG amplifier. Amplified EEG signals are recorded and stored using special BCI software. During an experiment, feedback is usually presented to the user either visually or acoustically.

After preprocessing, the next step in a BCI application is determining what characteristics of the signal are useful in representing the activity the subject is expected to perform. This is known as feature extraction. For example, in a motor imagery based BCI, a subject is required to imagine moving their left or right arm. While they perform this imagination, the power of the signal changes within a specific frequency region. In this case, the feature of interest is the spectral power of the EEG signal [23].

Over the course of an experiment, a subject typically performs some predetermined tasks repeatedly. By the end of the experiment, features corresponding to these different tasks can be collected and a classifier can be trained to distinguish between these tasks based on the features extracted.



## 2.2 EEG Signals

Electroencephalography (EEG) is a non-invasive method of measuring brain activity with the use of special electrodes made of Silver (Ag) or Silver Chloride (AgCl) [24]. A cap is first systematically placed on the subject's head. This cap has many holes on it representing the points where electrodes are placed and kept in place. Before placing these electrodes, a special gel (see Figure 2.4) is applied on the surface of the skin. This gel acts as a conductor between the electrode and the surface of the skin, and also helps stabilize the signal [25].



Figure 2.3: The electrodes used to record surface electrical activity of the brain.





Figure 2.4: Gel is first put on the surface of a subject's head before electrodes are placed. It is used to ensure the conductivity between the scalp and the electrodes.

In our work, electrodes have been placed according to the International 10-20 system [26], a proposition by the American EEG Society [20].

### 2.2.1 Sensorimotor Rhythms

13

located around the regions of the C3 and C4 electrodes.

Following this discovery, ERD and ERS have been used in a wide variety of applications over the following years. One of the earliest and groundbreaking applications of the ERD and ERS is motor imagery. In motor imagery, users are asked to imagine different kinds of movements such as left arm movements, right arm movements, and even right foot movements. Based on the corresponding EEG data, the BCI is able to classify what movement the user intends to do with a high degree of accuracy.

Another application of motor imagery is a study conducted by Wolpaw and his team in 1991 where motor imagery was successfully used to move a cursor to a target on a screen in a BCI system [29]. In this experiment, 8-12 Hz mu rhythm is used to move a cursor at the middle of a screen to a target located at either the top or the bottom of the screen. The feature used in this case is the amplitude of the mu rhythm. Larger mu amplitudes translated into an upward cursor movement and lower mu amplitudes translated into downward cursor movements. The mu amplitude is calculated by taking the square root of the power and then expressed in volts. This expression is compared to 5 different voltage ranges already predetermined by the operator. The result of this comparison produces 5 possible cursor movements, measured in a number of steps.

By 2004, Wolpaw et al. had succeeded in improving the system by accommodating two dimensional cursor control movements [30]. In this case, cursor movements depends on the result of a weighted linear combination of mu and beta amplitudes. To maximize performance, the weights are updated after every trial by using information obtained from trials that have already been performed.

Motor imagery applies to other parts of the body as well. For example, sensorimotor rhythms have also been shown to be applicable in a brain switch paradigm that involves foot motor imagery [31]. Similarly, a study has shown that it is possible to distinguish motor imagery related to the left hand, right hand, foot and tongue [32].

### **2.2.2 Event Related Potential**

Event Related Potentials (ERP) are positive or negative deflections in the EEG signal in response to certain psychological events [33]. ERPs can occur either before,

during, or after the psychological event. For similar events, ERPs usually occur at similar times [18].

In a study performed by Walter [34], it has been discovered that when subjects attempt to press a button right after seeing a target, negative deflections of large amplitudes can be observed moments before the subject actually press the button. This deflection is called the Contingent Negative Variation (CNV). CNV can be seen as an indicator that a subject is mentally preparing to execute a particular task. This is not to say that these deflections are easy to detect. In fact, the amplitudes of event related potentials can be very small in comparison to the background noise of the EEG signal. Because of this, multiple recordings of ERPs may need to be averaged if they are to be detected with a reasonable degree of accuracy.

Another well known ERP is called the P300, or P3 in short. It is a deflection that occurs 300ms after a subject experiences either a visual or auditory stimulus that that occur unexpectedly [34, 35, 36]. What is interesting about this signal is that the degree to which EEG deflects depends on how unexpected the stimulus is to the user. Following this discovery, several interfaces have been designed to enable users to use BCI by observing certain unpredictable stimuli and generating P300 signals as a result. This concept of paying attention while observing an unpredictable event is also known as the oddball paradigm [9].

In most P300-based BCI experiments, users are presented with a series of different visual, auditory, or haptic stimuli, each of which represents a different kind of output. By focusing on any one of these stimuli, users are able to generate a P300 signal in response to that stimulus, which can be detected by the BCI. Detecting P300 signals enables the BCI to make predictions about the user’s selection based on their brain activity.

One example of an oddball paradigm is the P300 speller initially developed by Donchin et al. [9]. The interface of a P300 speller is usually a 6 x 6 matrix consisting of letters, digits, or even symbols. Columns and rows of this matrix continuously flash in succession with the objective that the user selects a character from this matrix by simply staring at the letter and counting the number of times it flashes. Since the user

has no idea when their desired letter flashes, they generate a P300 signal whenever that letter flashes.

The intention of the user can be determined by analyzing the response generated by each flashing row and column. The row and column corresponding to the highest P300 component is selected and the product of their intersection is the letter selected by the classifier [37].

### 2.2.3 Error Related Potentials

Error related potentials are special kinds of ERPs that occurs when a user realizes a mistake in the output of a BCI that could be either as a result of the user or a misclassification on the part of the BCI. The first kind of ErrP occurring when a user performs an error and realizes it immediately is known as Response ErrP. The second type of ErrP that occurs as a result of a wrong output by the BCI is known as feedback ErrP [38].

Error related potentials typically have two deflection components. The first component is a negative deflection, also known as error related negativity (ERN), that occurs 100 ms after the user reacts to an error made either by themselves or by the BCI. This negative deflection occurs in the fronto central part of the brain. The second component is a positive deflection, also known as error positivity  $P_e$ , that occurs between 200ms and 500ms after the user realizes an error. This deflection occurs in the parietal region. All components of the ErrP can be recorded from the FCz, Cz, and Fz channels.

In a recent study by Chavarriaga et al. [1], it has been proposed that error related potentials could be used in learning an optimal decision making process in a classifier. This can be performed by decreasing the likelihood of the BCI repeating such a decision in the context that the error was performed. This idea is commonly known as Reinforcement Learning [39, 40]. Additionally, Chavarriaga and his colleagues tried to examine whether similar ErrPs are generated by users when they only observe the performance of the BCI via a monitor. In this study, users clearly have no control over the events presented on the BCI.

The protocol goes as follows; the user sits in front of a computer screen and monitors



the movement of a moving cursor, which is a green square. The cursor is designed to move towards a target, which is a colored square located at either the left or the right of the cursor. The objective is for the cursor to reach the target by moving in the desired direction in a series of steps. During these movements, there is a probability,  $P_{err}$ , that the cursor will move in the opposite direction and it is expected that when the user observes such a movement, they will produce an ErrP. Chavarriaga et al. were able to show that it was possible to detect the ErrP in such a scenario. The bold waveform in Figure 2.6 shows the overall average waveform of the error minus correct responses obtained in their work. It can be seen that the ErrP has two positive deflections at 200ms and 300ms and a negative deflection at 250 ms.

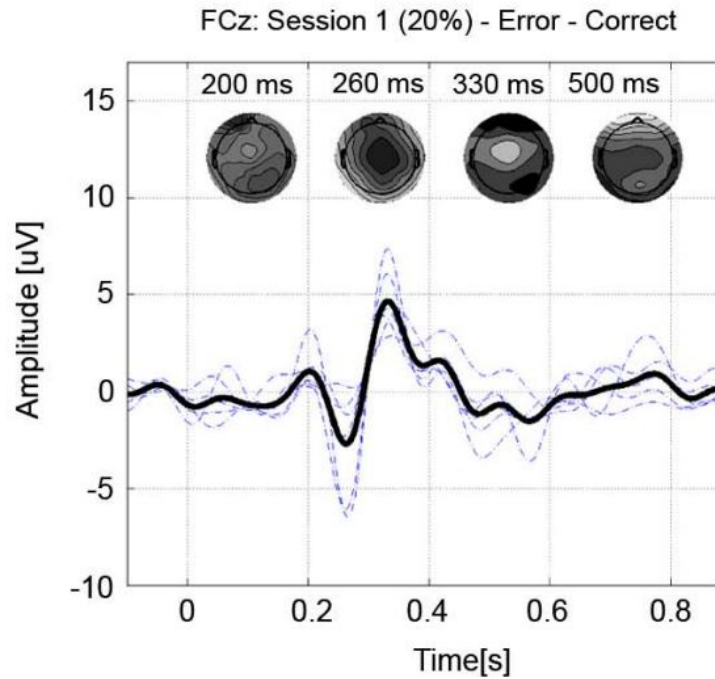


Figure 2.6: The overall average waveform of the error minus correct responses obtained from the work of Chavarriaga et al. [1].

In another work by Pierre W. Ferrez et al. [38], ErrP has also been successfully detected, but this time, rather than the user observing all events on a monitor, they have control of the movements. In their experiment, users pressed keys on a keyboard to move a robot towards a certain side of a room, which could be either to the left or

to the right. In this experiment, there is a certain probability that the robot moves in a direction that is opposite to the one intended by the user.

Recent work on ErrP includes a study by Iturrate et al. where latency correction of ErrP recorded from previous experiments has been shown to reduce BCI calibration time by 50% [41]. This is performed by computing a latency variation parameter,  $d_{E_i, E_j}$ , which is a shift that results in a maximum cross correlation between data collected from a previous experiment  $E_i$  and a current experiment  $E_j$ . Data collected from previous experiments are shifted according to this parameter and combined with a few samples recorded from a more recent experiment to train a classifier. Using this process has shown that BCI calibration time can be reduced significantly with data obtained from previous experiments.

ErrP has also been shown to be related to empathic attributes such as motivation, emotion, attention levels, and anxiety in a protocol that involves 2 steps [42]. First, empathic trait scores were obtained for each participant based on answers provided on a questionnaire. Second, the participants performed a gambling experiment to generate error related potentials. Results obtained in this study have shown a significant inverse correlation between ErrP amplitude and personal distress scores.

Error related potentials have also been detected in P300 based BCI systems as well. In 2010, Bernardo et al. attempted an online detection of ErrP in P300 spellers with roughly 60% accuracy [43]. In their experiment, they first implemented a P300 speller that used a generic algorithm in order to detect the P300 signals, and then included an automatic error-correction system that is based on the detection of ErrPs from a single sweep of data. In their case, it is shown that it is in fact possible to not only detect ErrP, but that it is also theoretically possible to use that information to improve the performance of a P300 speller. A similar piece of work has also shown that an error correction system based on ErrPs can increase information transfer rates in P300 based BCIs for both healthy and motor impaired subjects [44].

Zeyl et al. have shown that ErrP and P300 scores can be used in P300 based BCIs simultaneously. In their first work, ErrP and P300 scores generated by a bidirectional stepwise linear discriminant analysis [45] are fed into a random forest error detector

with 100 decision trees [46]. Results of this work indicate that using P300 scores alone provides better classifier adaptation in comparison to results obtained when ErrP and P300 scores are combined. However, one year later, another study from the same group showed that ErrP scores can provide better results when combined with P300 scores obtained from a real-time Bayesian dynamic framework [47]. Bayesian dynamic stopping is a mechanism where rows and columns in a P300 speller repeatedly flash until a confidence criterion is reached and it has already been shown to improve performance in P300 spellers [48]. This approach builds on Zeyl’s previous work by combining P300 scores generated through this Bayesian dynamic stopping framework and ErrP scores generated by a random forest error detector as in the previous study. Results have shown an improvement in the speed and accuracy of the system.

Variation in error related potentials have also been shown to be affected by attributes such as intolerance of uncertainty [49]. Intolerance of uncertainty is a natural predisposition to feel threatened by uncertain events. Other works in this context have also shown that ErrP is more pronounced in subjects who are more emotionally responsive to errors such as those with anxiety disorder [50], obsessive compulsive disorder [51], and pathological worry [52].

## 2.3 Classification Methods

### 2.3.1 Linear Discriminant Analysis

Linear Discriminant Analysis is a classification method that uses hyperplanes constructed from a linear combination of the features of each class to separate the classes. It is assumed that the classes are represented by a normal distribution. The aim is to construct a feature  $y$  that is a linear combination of the data  $x$  and effectively compress all classification related information into one feature. This can be done by finding a plane where the two classes are separated the most. This decision boundary can be written as

$$y = w_1x + w_0 \tag{2.1}$$

such that any data point  $x$  satisfying  $y > 0$  is classified into say Class 1, and that which satisfies  $y < 0$  is classified into Class 2.

### 2.3.2 Support Vector Machines

Support Vector Machines is quite similar to LDA in the sense that the aim is to separate classes by using hyperplanes. The main difference between SVM and LDA is that while LDA generates the hyperplane by using posterior probabilities, SVM generates a hyperplane that maximizes the distance between the hyperplane and the data points. Hence, SVM can also be referred to as a Maximum Margin Classifier. If the hyperplane is represented as shown in equation 2.1, then finding such a hyperplane is equivalent to solving the following optimization problem [53].

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \mathcal{J}(w, w_0) = \frac{1}{2} \|w\|^2 \\ \text{subject to} \quad & y_i(w^T x_i + w_0) \geq 1, \quad i = 1, \dots, N. \end{aligned} \tag{2.2}$$

where  $y_i$  is an indicator function such that  $y_i = +1$  for Class 1 and  $y_i = -1$  for Class 2.

### 2.3.3 Gaussian Mixture Models

The Gaussian distribution is arguably the most common probability distribution used in BCI applications. Despite its many analytical properties, it could have shortcomings when it comes to real datasets [54]. That is not to say that Gaussian distributions are not useful in such cases. In fact, by using a linear superposition of two or more Gaussian distributions, it can be possible to more accurately capture the statistical properties of some real datasets. Mathematically, the superposition of  $N_p$  Gaussian densities, or a mixture of Gaussians, can be written in the following form.

$$p(x) = \sum_{k=1}^{N_p} \pi_k \mathcal{N}(x | \mu_k, \Sigma_k) \tag{2.3}$$

Each Gaussian distribution comprising the mixture is known as a component with each component having its own mean  $\mu_k$ , covariance  $\Sigma_k$ , and mixing coefficient  $\pi_k$ . The

mixing coefficients determine the weight of each Gaussian mixture and are normalized such that,

$$\sum_{k=1}^{N_p} \pi_k = 1 \quad (2.4)$$

In this thesis, it is assumed that the prototypes to be used have equal mixing coefficients such that

$$p(x) = \frac{1}{N_p} \sum_{k=1}^{N_p} \mathcal{N}(x|\mu_k, \Sigma_k) \quad (2.5)$$

If each class  $k$  is modeled by using  $N_p$  components, then the activity  $a_k^i$  of the  $i$ th prototype of class  $C_k$  for a specific data point  $x$  can be written as,

$$a_k^i(x) = |\Sigma_k|^{-\frac{1}{2}} \exp(-\frac{1}{2}(x - \mu_k^i)^T \Sigma_k^{-1} (x - \mu_k^i)) \quad (2.6)$$

where the constant terms have been dropped for convenience. This approach is the same as that used by [55]. Based on this, the posterior probability of each class  $C_k$  becomes,

$$y_k^i(x) = p(x|C_k) = \frac{a_k^i(x)}{A(x)} = \frac{\sum_{i=1}^{N_p} a_k^i(x)}{\sum_{k=1}^K \sum_{i=1}^{N_p} a_k^i(x)} \quad (2.7)$$

where  $a_k^i(x)$  is the activity of class  $C_k$  and  $A(x)$  represents the total activity of the network.  $y_k^i(x)$  can be viewed as the *responsibility* that class  $C_k$  takes in explaining the data  $x$ . The class producing the highest level of activity for any given data is selected as the response of the classifier. In this work, the covariance matrices are assumed to be diagonal.

The means of the Gaussian components will have to be initialized before the classifier can be trained. This can be achieved with the help of a clustering algorithm,  $k$ -means for example, after which the covariance matrix can also be initialized in the following form.

$$\Sigma_k = \frac{1}{|S_k|} \sum_{x \in S_k} (x - \mu_k^{i*})(x - \mu_k^{i*})^T \quad (2.8)$$

where  $S_k$  represents the set of samples that belong to class  $C_k$ ,  $|S_k|$  represents the cardinality of  $S_k$ , and  $i^*$  represents the prototype that is closest in distance to  $x$ .

During training, the estimates of  $\mu_k$  and  $\Sigma_k$  are improved for every data sample trained on the classifier. This is performed through a stochastic gradient descent algorithm that minimizes the mean square error of the class posterior probability,  $E = \Sigma_k(y_k - t_k)^2$ , where  $t_k$  is the target vector for class  $C_k$ . The target vector takes a form of 1-of-c, *i.e.*, the target vector for the *correct class* is  $(1, 0)$  and that of the error class is  $(0, 1)$ . The gradient of the mean square error is then,

$$\Delta\mu_k^i(x) = \alpha \frac{\partial E}{\partial \mu_k^i}(x) = \alpha \frac{a_k^i(x)}{A(x)} \frac{(x - \mu_k^i)}{\Sigma_k} e_k(x) \quad (2.9)$$

and then,

$$\Delta\Sigma_k^i(x) = \beta \frac{\partial E}{\partial \Sigma_k^i}(x) = \beta \frac{a_k^i(x)}{A(x)} \frac{(x - \mu_k^i)^2}{(\Sigma_k)^3} e_k(x) \quad (2.10)$$

$$e_k(x) = (t_k(x) - y_k(x)) - \sum_j y_j(x)(t_j(x) - y_j(x)) \quad (2.11)$$

Here,  $\alpha$  and  $\beta$  are the learning rates. When all means and covariance matrices are updated, the covariance matrices of all prototypes for each class are averaged, resulting in the common-class covariance matrix. This helps improve the performance and robustness of the classifier.

## 2.4 Adaptation in Brain Computer Interfaces

### 2.4.1 Nonstationarity in EEG

Statistical properties of EEG signals change over time [56]. This nonstationarity has been studied in detail by Kaplan et al. [57] and it has also been shown to be deterrent to BCI performance over the course of an experiment [58]. This is because a classifier trained optimally at any given time becomes suboptimal when the conditions under which it is trained have changed.

Kaplan et al. emphasized that the signal processing techniques used to filter and extract information from EEG rely on one basic assumption: the assumption that EEG signals are stationary. However, that is not the case. In most cases, nonstationarities can be eliminated or even ignored. This can be done by smoothing or averaging the signal. This approach makes data from relevant electrodes more distinct from their neighboring electrodes. This procedure might result in loss of useful information. In some sense, one fundamental question is how to reliably maintain the optimality of a classifier during an experiment even when the statistical nature of the EEG signals change.

One answer to this question is adaptation in BCI systems. Adaptation seeks to reconcile the issue of an ever-changing EEG signal by adapting the classifier to these changes based on new information.

## **2.4.2 P300 Based BCI**

Section 2.2.2 introduced the P300 speller that enables users to type letters using EEG. Despite research efforts that yield high performances, there is still room for improvement. One major challenge researchers face is the high number of flashes required before a confident selection can be made. This makes the decision making process very long and tiring over the course of an experiment. For example, in one study, each character had been flashed 15 times [59].

Various studies aimed at adapting P300 based BCIs have been done. Most of these studies fall into two categories of adaptation:

1. Adaptation based on statistical properties of EEG [60, 61, 62].
2. Adaptation based on new information, usually in the form of a feedback, such as the use of error related potentials [63, 64, 65].

By tracking statistical properties of EEG and identifying the changes that occur over time, it is in principle possible to update a classifier accordingly. A study by Y. Li et al. for example, has shown that a self-training SVM classifier can not only improve performance in P300 BCIs, but it can also reduce the effort required to train

the classifier [66]. In another study, two classifiers - BLDA and FLDA - have been adapted to nonstationarities in EEG data simultaneously in a collaborative adaptation process [67]. The basic idea in this co-training approach is that new data are classified by both classifiers. The labels produced by these two classifiers may not necessarily be identical and to take advantage of this, each classifier is updated by using the label produced by the other classifier, rather than its own. This in principle enables both classifiers to learn not only from EEG but also from each other. This approach has shown promising results in comparison with other approaches used in the literature.

Adaptation based on new information is most commonly seen in the form of error related potentials. This approach attracts researchers because it addresses some core issues posed by the adaptation process that is based on statistical properties of EEG. Adaptation can be performed by using the labels produced by the classifier from the test data and this has been shown to have an improved performance when compared to a system where no such information is used [68, 69]. However, such a label may be incorrect and adapting the classifier with wrong information is undesirable. ErrP detection is attractive because in case of perfect detection, the wrong labels may easily be identified.

In a recent study, ErrP has been detected online in the context of a P300 speller [64]. In their setup, a currently selected letter is ignored if an error related potential is detected. Otherwise, it is assumed to be true and the experiment proceeds. Training the ErrP classifier involved the same setup with a slight difference when it comes to the feedback. During training, 20% of the feedback provided is incorrect. This in effect, implies that 20% of the data contains error related potentials. The results of this study indicated that it is possible to detect error related potentials online in a P300 speller paradigm.

Two years later, this was taken one step further [65]. In this work, ErrP detection is used to make an educated guess on what the actual intention might be. The experimental setup in this work incorporated an online correction procedure that replaces a current selection with the second best guess of the classifier in the event that an ErrP has been detected. Their study has been able to show that as long as a subject



remains engaged in the experiment, online correction can improve the performance of P300 spellers.

The focus of this thesis is an investigation of the detectability of ErrP in different BCI protocols. In particular, we wanted to understand how ErrP detection is affected in three different P300 speller protocols.

1. When a subject merely observes the interface and recognizes when the system makes an error.
2. When a subject actively determines the letters to be selected by pressing keys on a keyboard, and then observing whether the displayed letter was in fact the one they had selected or not.
3. When a subject actively uses EEG to spell letters during the experiment.

### **2.4.3 Motor Imagery Based BCI**

There have been various studies on motor imagery based BCI systems. Even though these studies report quite high accuracy rates for motor imagery classification, it is believed that there is still room for improvement. In 2004, Vidaurre et al. worked on designing a Quadratic Discriminant Analysis (QDA) classifier whose covariance matrices are updated after every trial [68]. Each experiment starts with an initial classifier that has been trained with data collected from over 1620 trials. After every trial, the information matrix of the classifier is updated with a coefficient determined by an optimization process. Their work was able to show that the information matrix did indeed change over time and that an improvement in performance is possible with their adaptation process in motor imagery based BCI.

Four years later, Vidaurre et al. introduced co-adaptive learning, a concept that is one step ahead of adaptive learning [70]. In contrast to adaptation where a classifier changes with respect to the changes in the user's EEG, co-adaptation enables the subject to learn from the performance of the classifier and adapt to it as well. In this case, the learning capacity of a user is acknowledged and the classifier is designed in order to

benefit from their capacity. This can be achieved by splitting the adaptation process into three levels, each level accompanied by an increased dependency on the user's EEG. The results of this experiment showed that 10 out of 14 users have been able to improve their BCI performance and that the number of users previously unable to control a BCI decreased from 25% to 12.5%.

Other ways of adaptation in motor imagery based BCI have also been investigated. For example, McFarland and his colleagues were able to show that two adaptation procedures, adaptation of feature weights and adaptive normalization, are able to improve BCI performance [71]. Feature weights were assigned based on a predetermined training set, which could either be the first session, the current session, the preceding session, or all sessions combined. Adaptive normalization involved normalizing the EEG data before feature extraction with the aim of reducing the effect of nonstationarity. Their work was able to show that for these two techniques, BCI performance could be improved.

The essence of adaptation is primarily the use of new information in order to help make better decisions by designed classifiers. Error related potentials become relevant because they provide new information that can be useful in the adaptation process. Ricardo Chavarriaga and his colleagues detected ErrP signals during a BCI experiment in 2010 [1]. Their work reports the discovery of ErrP waveforms consistent with those reported by previous studies. They have also showed that classification performance of ErrP was dependent on the frequency of errors in the BCI.

In another study, Ferrez et al. investigated the possibility of simultaneously detecting ErrP while classifying motor imagery in real time [72]. The protocol used to perform this task provided a one second interval that was used for motor imagery classification and then a 400ms window to be used in detecting the presence of error related potentials. If ErrPs were detected, the intended movement would be canceled, else nothing would be done. Their results showed that some subjects were able to achieve satisfactory motor imagery performance while maintaining a high level of ErrP detectability.

In this thesis, we analyzed the detectability of error related potentials within the

context of a motor imagery BCI. Given that various studies have shown that the nature of the ErrP could change depending on context, we wanted to understand these differences. We wanted to investigate how a change in context affects ErrP detectability. For this, we used three different protocols.

1. When a subject merely observes the interface and recognizes when the system makes an error.
2. When a subject actively presses a key to initiate the movement of a ball on the screen, and then observing whether the ball in fact did move in the desired direction or not.
3. When a subject actively uses EEG to move a ball using motor imagery in an experiment.

## Chapter 3

# Analysis of ErrP in P300 based Brain Computer Interfaces

This chapter focuses on investigating the nature and detectability of error related potentials across three different types of P300 based BCI protocols. For simplicity, give name these protocols as follows.

1. P300 Observe: The BCI automatically spells out words while the user monitors its performance.
2. P300 Control: The subject is required to type words using the keyboard while monitoring the response of the computer when the key is pressed.
3. P300 EEG: The subject uses EEG to spell words during an experiment.

One interesting property of these three protocols is that they can also be seen as three different forms of user engagement. In some sense, one aim of this study is to see whether user engagement type can be used as a basis for explaining the results obtained in these experiments.

The P300 system used in this work was originally developed by Amcalar [25] in the BCI Lab of Sabancı University. It features a 6x6 matrix filled with 26 letters, digits 1 to 9, and an underscore representing the space bar key. This interface has been modified in three different ways corresponding to the three P300 based BCI protocols.

### 3.1 Experimental Description

We analyze error related potentials in P300-based BCIs using a setup similar the setup used by Yilmaz [69]. It is an updated version of the SU-BCI P300 stimulus interface with the added capacity of detecting error related potentials if and when they exist. The interface starts with a main page (see Figure 3.1) where subject-specific and experiment-specific information are entered.

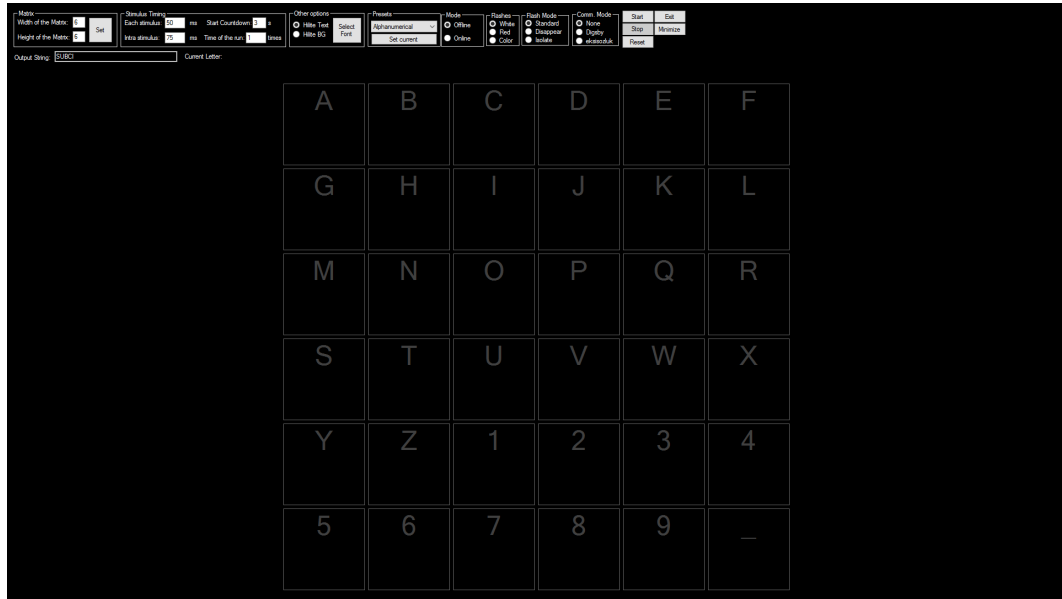


Figure 3.1: The main page of the P300 related experiments where experiment-specific settings are entered.

Each of the three different protocols have a similar structure; a target phase, a selection phase, and a feedback phase.

1. A target phase, where the letter to be typed is shown in grey color. This gives the subject an opportunity to know where the letter is located before flashing begins. The target phase lasts for one second at the beginning of every trial and is identical for all protocols.

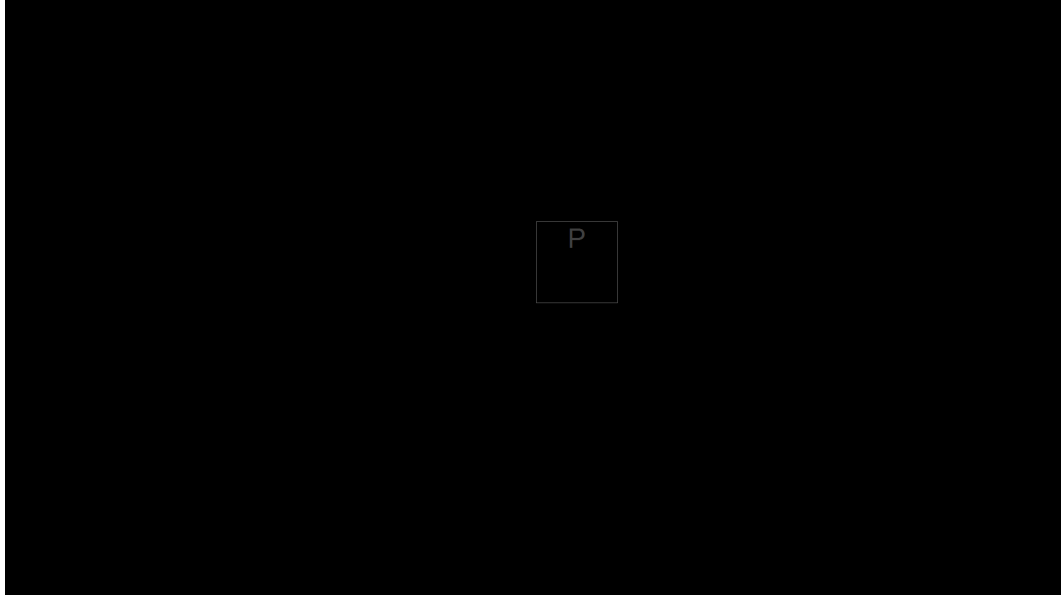


Figure 3.2: Target Phase of P300 based BCI experiments. Subjects are shown a random letter for one second.

2. A selection phase, when a letter is selected by the BCI as rows and columns flash randomly on the screen. This phase differs between all three protocols and these differences shall be discussed in the corresponding sections.

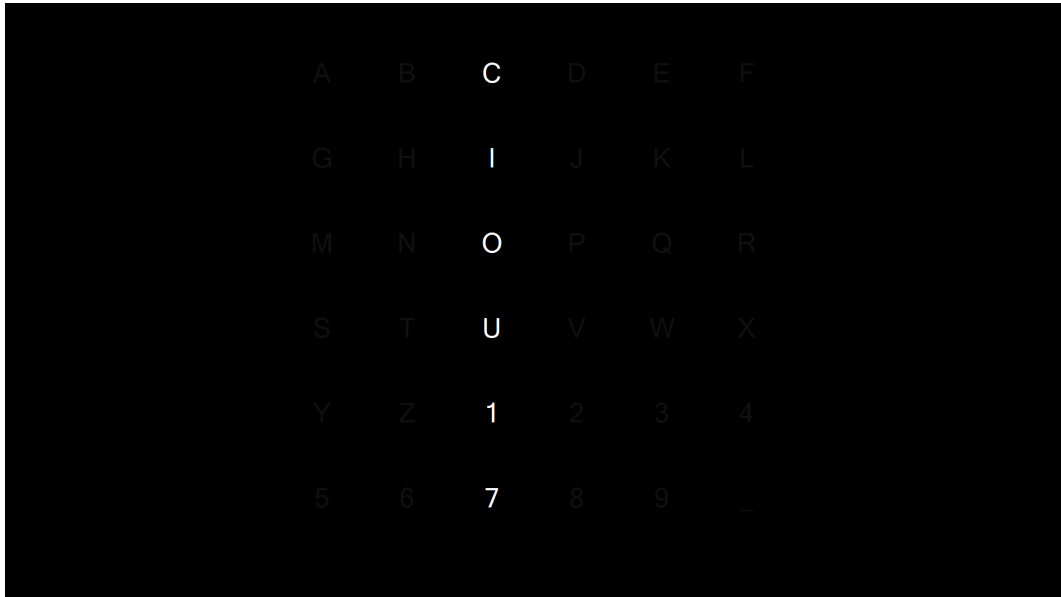


Figure 3.3: Selection Phase of P300 based BCI experiments. Rows and columns randomly flash during this phase.

3. A feedback phase, where the letter selected by the BCI is presented in green color. This phase is identical for all three protocols as well.

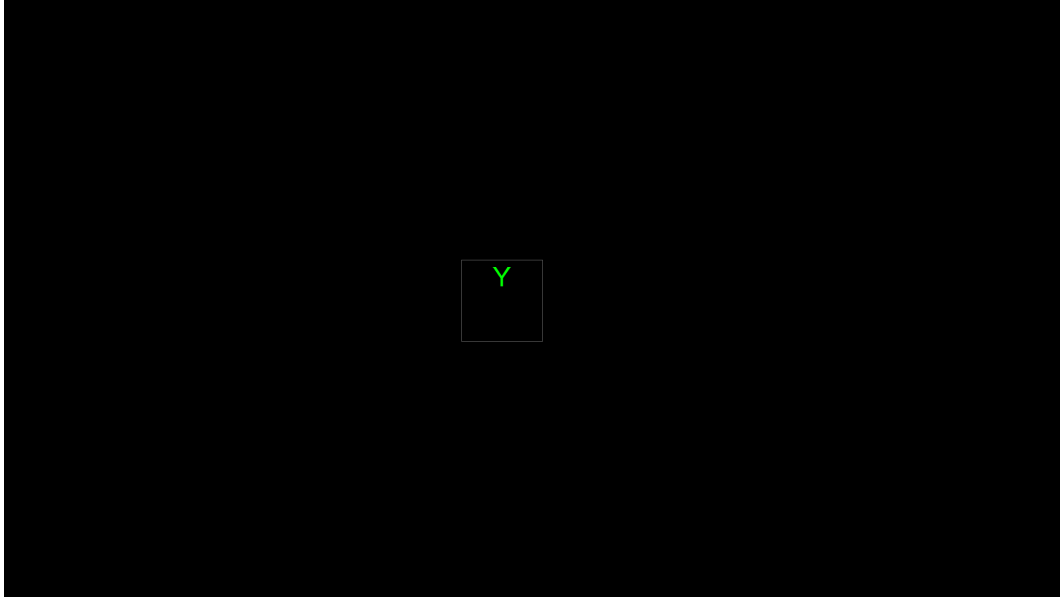


Figure 3.4: Feedback Phase of P300 based BCI experiments. A feedback is presented to the subjects and they are instructed to notice if it is the same as the letter presented during the target phase.

Each protocol is divided into 3 runs, each containing 100 trials (letters to be typed) with a 5 minute rest between each run. The characters typed during each run consist of a repetition of the sentence “THE\_QUICK\_BROWN\_FOX\_JUMPS\_OVER\_THE\_LAZY\_DOG\_” until 100 characters are exhausted. During experiments, the feedback is always programmed to be 70% correct.

### 3.1.1 Data Processing

ErrP has been shown to be detectable with EEG recorded from the FCz and Cz electrodes [73, 74, 75] so we use these electrodes in our experiments as well. We record EEG from 8 different electrodes, Fz, FCz, FC1, FC2, Cz, C1, C2, and CPz, sampled at 2048 Hz. Electrodes apart from FCz and Cz are used as referencing electrodes. We reference each electrode by subtracting from it the mean of data collected from the four adjacent electrodes, that is, top, bottom, left, and right. Following this, data from the

FCz and Cz electrodes are downsampled to 64 Hz and then passed through a band pass filter with cutoff frequencies 1-10Hz. The epochs corresponding to the feedback phase are extracted and categorized into two; error and correct samples. Each epoch contains 64 samples and is used as feature for the classification process.

### **3.1.2 Classification**

The first two runs of each experiment are used as a training set and the last run as a testing set. The classifier trained on the training set is now tested on the testing set and this process is performed 100 times to obtain robust results. In this thesis, four different classifiers have been investigated; SVM, LDA, Decision Tree, and a Gaussian Mixture Model classifier.

## **3.2 P300 Observe**

### **3.2.1 Experimental Description**

This experiment aims to understand the detectability and nature of error related potentials generated in the context of a P300 based BCI when a subject only observes the performance of the BCI. During the selection phase, the subject is instructed to do nothing. This is ensured by flashing each row and column only once for a period of 25 ms with an interstimulus interval of 25 ms. In a practical P300 based BCI, these time intervals are too small. However, it is not relevant in this piece of work because P300 detection is not considered.



### 3.2.2 Results

Table 3.1 provides ErrP classification results for four different classifiers – SVM, LDA, Decision Tree, and GMM – for the *Observe* protocol of the P300 based BCI experiments. Results are provided for FCz and Cz electrodes for all subjects. For each electrode, correct classification percentage is provided for both correct and error samples.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	69.71	64.88	69.45	54.00
	Subject 2	69.64	66.44	70.28	55.00
	Subject 3	73.88	60.05	75.50	63.66
	Subject 4	73.75	64.10	61.85	54.22
	Subject 5	75.19	66.15	76.17	67.36
LDA	Subject 1	73.30	53.77	73.07	55.22
	Subject 2	75.40	56.44	73.92	57.00
	Subject 3	71.04	54.77	75.26	57.72
	Subject 4	73.87	58.63	72.82	54.52
	Subject 5	75.14	62.84	75.46	60.78
Decision Tree	Subject 1	76.16	65.72	74.14	64.11
	Subject 2	75.85	63.61	75.42	64.61
	Subject 3	73.88	64.44	75.02	64.66
	Subject 4	74.31	65.31	74.29	64.00
	Subject 5	74.53	64.52	75.39	65.00
GMM	Subject 1	69.64	63.66	68.09	62.16
	Subject 2	69.71	63.83	68.80	63.72
	Subject 3	65.45	59.94	70.85	64.22
	Subject 4	65.73	60.15	59.17	51.47
	Subject 5	68.29	63.21	70.12	65.89

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)

Table 3.1: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the *Observe* protocol in P300 based BCI experiments. For each electrode, the accuracy of classifying correct and error samples are presented.

## 3.3 P300 Control

### 3.3.1 Experimental Description

This experiment investigates the nature and detectability of error related potentials in the context of a P300 based BCI when a subject selects a letter by pressing keys on a keyboard. During the selection phase, the subject is instructed to press the letter they have seen in the target phase on the keyboard. In addition to this, no flashes occur in the selection phase of this experiment.

### 3.3.2 Results

Table 3.2 provides ErrP classification results for four different classifiers; SVM, LDA, Decision Tree, and GMM in the *Control* protocol of the P300 based BCI experiments. Results include classification performance for both correct and error samples for FCz and Cz electrodes.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	75.19	64.89	72.80	60.84
	Subject 2	73.07	61.89	75.48	63.94
	Subject 3	73.69	64.38	74.33	65.90
	Subject 4	70.00	59.52	69.76	66.47
	Subject 5	73.37	63.40	74.35	65.65

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
LDA	Subject 1	74.82	60.15	74.04	58.63
	Subject 2	72.95	58.36	73.78	58.47
	Subject 3	73.05	60.42	70.07	58.66
	Subject 4	72.58	58.09	72.33	58.19
	Subject 5	72.10	58.50	73.65	59.45
Decision Tree	Subject 1	75.34	67.10	76.43	65.42
	Subject 2	74.24	65.21	74.07	66.47
	Subject 3	75.51	69.81	75.10	66.14
	Subject 4	74.20	67.28	73.56	67.52
	Subject 5	73.85	67.60	76.22	66.40
GMM	Subject 1	63.60	63.48	62.10	61.48
	Subject 2	64.39	61.30	59.62	59.62
	Subject 3	68.28	61.84	65.05	63.20
	Subject 4	66.61	61.87	70.76	59.38
	Subject 5	64.43	59.82	68.75	63.75

Table 3.2: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the *Control* protocol in P300 based BCI experiments. For each electrode, the accuracy of classifying correct and error samples are presented.

### 3.4 P300 EEG

In this protocol, the selection phase is designed such that each row and column flashes twice with each flash lasting 25 ms. The interstimulus interval is also set to 25 ms. During this phase, the subject is instructed to count the number of times the target letter flashes and is told that the BCI determines the target letter based on their EEG signals. In fact, the letter outputs are generated randomly with a specified error probability of 30%. Following this, the subject is instructed to observe whether the

letter presented during the feedback phase is the same as the target letter. We evaluate ErrP classification performance on four different classifiers; SVM, LDA, Decision Tree, and GMM.

### 3.4.1 Results

ErrP classification results for the *EEG* protocol of the P300 based BCI experiments are presented in Table 3.3. Results are provided for four different classifiers for all subjects. Additionally, classification performances are calculated for both correct and error samples over FCz and Cz electrodes.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	79.87	71.58	74.43	62.88
	Subject 2	80.00	72.56	79.51	69.37
	Subject 3	79.85	71.65	74.39	65.10
	Subject 4	74.78	65.25	71.19	64.65
	Subject 5	77.56	69.05	75.56	68.40
LDA	Subject 1	78.02	66.41	74.36	60.76
	Subject 2	79.02	64.68	77.34	60.37
	Subject 3	77.46	68.10	74.82	60.55
	Subject 4	75.34	61.00	74.39	63.35
	Subject 5	78.02	63.65	74.56	61.55
Decision Tree	Subject 1	78.48	70.11	76.31	68.00
	Subject 2	77.19	70.06	76.36	66.75
	Subject 3	78.17	70.75	73.95	67.60
	Subject 4	76.19	68.80	74.82	65.80
	Subject 5	74.31	65.40	75.46	67.60

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
GMM	Subject 1	76.46	69.41	71.43	67.05
	Subject 2	73.00	64.68	70.65	66.06
	Subject 3	74.65	65.55	71.34	65.11
	Subject 4	66.65	64.15	63.26	62.30
	Subject 5	69.70	63.65	69.26	66.10

Table 3.3: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the *EEG* protocol in P300 based BCI experiments. For each electrode, the accuracy of classifying correct and error samples are presented.

## 3.5 Discussion

### 3.5.1 Waveform analysis

Figure 3.5 shows average waveforms computed for FCz and Cz electrodes for all three protocols. Blue lines correspond to the average correct waveform across all subjects. Red lines correspond to the average error waveforms across all subjects. Yellow waveforms are the overall error-minus-correct waveforms computed by taking the difference between the blue and red waveforms. The first column represents waveforms calculated from the FCz electrode and the second column represents that from the Cz electrode. From top to bottom, the rows correspond to results obtained for the *Observe*, *Control*, and *EEG* protocols.

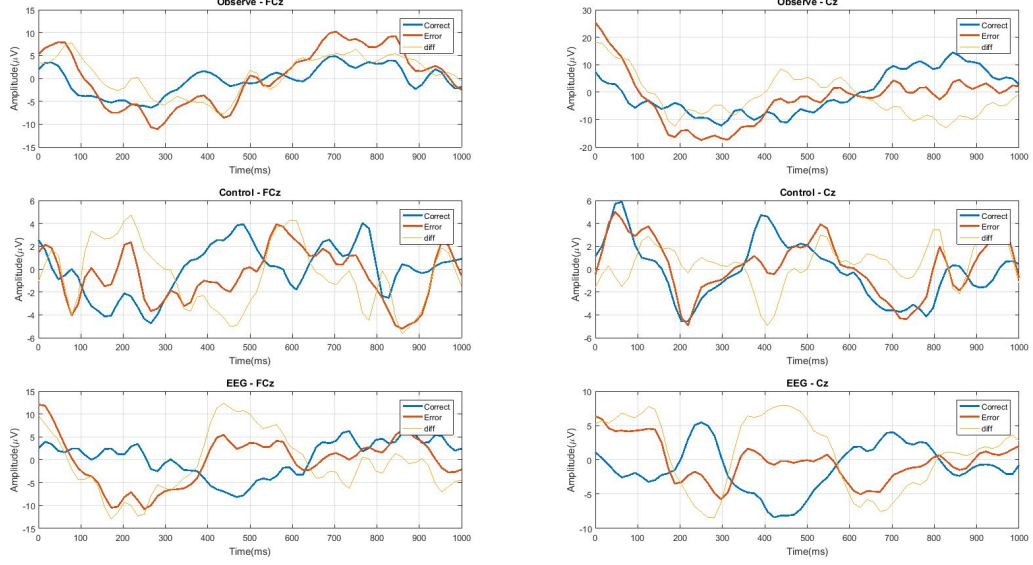


Figure 3.5: Average waveforms over all subjects for correct (blue), error (red) and error-minus-correct (yellow) samples for all three protocols. First column represents results obtained from the FCz electrode and the second column represents results obtained from the Cz electrode. Top row represents results for the *Observe* protocol, while the middle row represents the *Control* protocol and the bottom row represents the *EEG* protocol.

Figure 3.6 shows the mean error minus correct waveforms for all three protocols conducted. We have observed that the waveforms of all three protocols look similar and are consistent with the waveform obtained by Schmidt et. al. [76]. The similarities between the three waveforms were tested by first aligning the signals using a cross correlation procedure and then the correlation between each pairwise combination of error minus correct waveforms presented in Figure 3.6 is computed. The correlation coefficient between the waveforms of *EEG* and *Observe* is  $\rho = 0.26$ . For the waveforms of *EEG* and *Control*, the correlation coefficient is  $\rho = 0.37$  and the correlation coefficient between the waveforms of *Observe* and *Control* is  $\rho = 0.62$ .

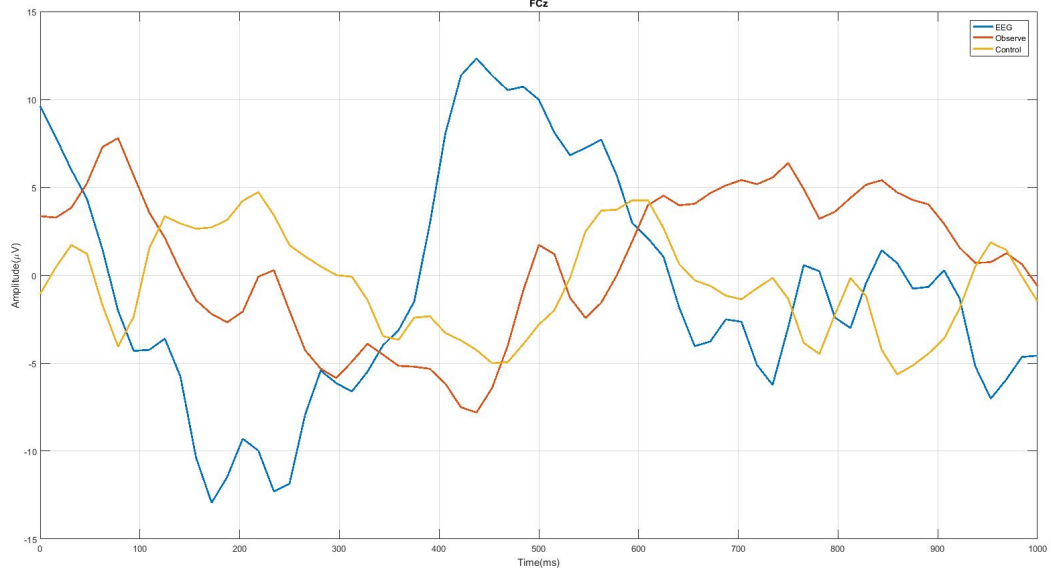


Figure 3.6: This figure shows the mean error minus correct waveforms for all three protocols as recorded from the FCz electrode. Blue represents *EEG* protocol, red represents *Observe* protocol and yellow represents *Control* protocol.

ErrPs generated in these P300 based protocols seem to have different waveforms compared to the motor imagery based BCI experiments described in the literature. They exhibit a wide negative deflection followed by a wide positive deflection. For the *EEG* protocol, the first negative deflection starts at around 200ms and the positive deflection at 450ms which is different from other protocols. Negative deflection starts at 450ms and ends at 550ms for *Observe* and it is from 500ms to 600ms for *Control*. Similar waveforms have also been reported by Spüler et. al. [77]. Latency analysis shows that waveforms produced by *Control* and *Observe* have a phase difference of 50 ms which is statistically significant with p-values much less than 0.01. A statistically significant phase difference has been observed between the waveforms of *EEG* and *Observe* protocols with very low p-values as well. The mean latency between the waveforms of *EEG* and *Observe* has been observed to be 300ms. The phase difference between *EEG* and *Control* has also been observed to be statistically significant with very low p-values and a mean difference of 250ms.

Experiment Type	Max. Negative Deflection( $\mu V$ )	Max. Positive Deflection( $\mu V$ )
<i>Observe</i>	-9.19 ( $\pm 2.09$ )	10.19( $\pm 1.84$ )
<i>Control</i>	-7.73 ( $\pm 1.58$ )	6.60( $\pm 1.50$ )
<i>EEG</i>	-13.63 ( $\pm 1.81$ )	13.25 ( $\pm 2.22$ )

Table 3.4: Maximum positive and negative deflections for the average error-minus-correct waveforms obtained for all three protocols.

### 3.5.2 Performance analysis

Results have shown that error related potentials can be detected with accuracies above 50% in all protocols and for all classifiers. In addition to this, we sought to further understand the performance of each classifier in comparison with the other classifiers used. For this to be relevant, we decided to treat this problem as if it were an online classification problem by doing two things. First, we select the electrode with the best performance in the training session for each subject. Second, we use the classification results of the selected electrode during the testing session as the actual performance of the classifier. This, we believe, gives a better reflection of what would have happened if such an experiment were performed online. The results of this procedure can be seen in Figure 3.7.



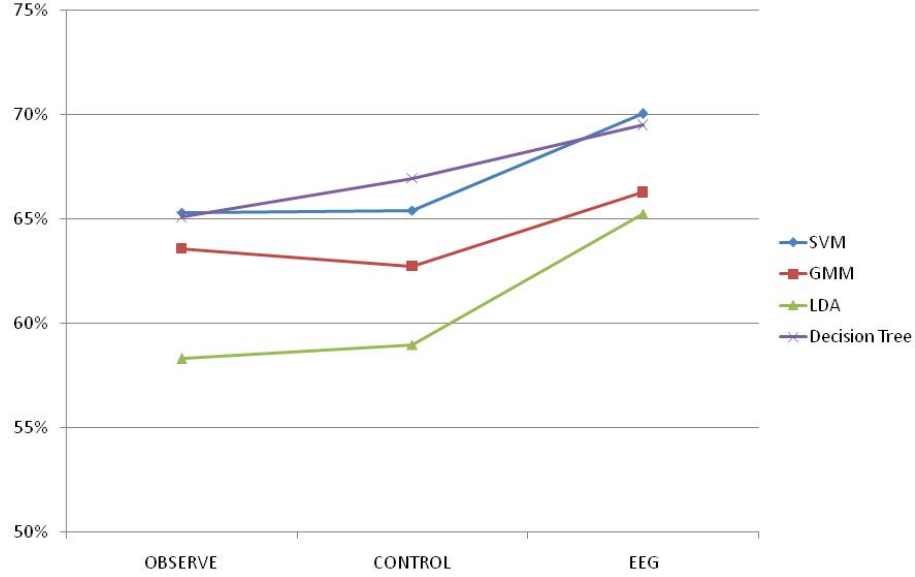


Figure 3.7: This figure presents mean accuracies obtained over all subjects and in all three protocols for four different classifiers; SVM, LDA, Decision Tree and GMM.

Figure 3.7 shows the mean classification results across all subjects for all four classifiers used. Decision Tree and SVM classifiers perform best and LDA performs with the lowest accuracy. GMM classifier also seems to perform relatively better compared to the LDA classifier. Indeed, a t-test performed on the mean accuracies between all classifiers indicate that the difference between the performances of the SVM and Decision Tree classifier is not significant ( $p = 0.92$  for *EEG*,  $p = 0.44$  for *Observe*, and  $p = 0.10$  for *Control*). The difference between the performances of the LDA and GMM classifier is not significant for *EEG* ( $p = 0.27$ ) but significant for *Observe* ( $p < 0.01$ ) and *Control* ( $p < 0.01$ ). All other combinations provide a significant difference with very low p-values.

Across experiments, Table 3.5 provides the results of a MANOVA test on the average classifier performances of each pairwise combination of the three protocols.

Protocols \ Classifiers	Classifiers			
	SVM	LDA	Decision Tree	GMM
<i>Observe, Control</i>	$<< 0.01$	0.27	$<< 0.01$	0.11
<i>Observe, EEG</i>	$<< 0.01$	$<< 0.01$	$<< 0.01$	0.09
<i>EEG, Control</i>	$<< 0.01$	$<< 0.01$	$<< 0.01$	$<< 0.01$

Table 3.5: MANOVA test results for pairwise combinations of the three protocols in the P300 based experiments.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	69	64
	Subject 2	FCz	69	66
	Subject 3	Cz	75	63
	Subject 4	FCz	73	64
	Subject 5	Cz	76	67
LDA	Subject 1	Cz	73	55
	Subject 2	Cz	73	57
	Subject 3	Cz	75	57
	Subject 4	FCz	73	58
	Subject 5	FCz	75	62
Decision Tree	Subject 1	FCz	76	65
	Subject 2	Cz	75	64
	Subject 3	Cz	75	64
	Subject 4	FCz	74	65
	Subject 5	Cz	75	65

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
GMM	Subject 1	FCz	69	63
	Subject 2	FCz	69	63
	Subject 3	Cz	70	64
	Subject 4	FCz	65	60
	Subject 5	Cz	70	65

Table 3.6: Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the *Observe* protocol of P300 based BCIs.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	75	64
	Subject 2	Cz	75	63
	Subject 3	Cz	74	65
	Subject 4	Cz	69	66
	Subject 5	Cz	74	65
LDA	Subject 1	FCz	74	60
	Subject 2	Cz	73	58
	Subject 3	Cz	70	58
	Subject 4	FCz	72	58
	Subject 5	Cz	73	59
Decision Tree	Subject 1	FCz	75	67
	Subject 2	Cz	74	66
	Subject 3	Cz	75	66
	Subject 4	FCz	74	67
	Subject 5	FCz	73	67

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
GMM	Subject 1	FCz	63	63
	Subject 2	FCz	64	61
	Subject 3	Cz	65	63
	Subject 4	FCz	66	61
	Subject 5	Cz	68	63

Table 3.7: Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the *Control* protocol of P300 based BCIs.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	79	71
	Subject 2	FCz	80	72
	Subject 3	FCz	79	71
	Subject 4	FCz	74	65
	Subject 5	FCz	77	69
LDA	Subject 1	FCz	78	66
	Subject 2	FCz	79	64
	Subject 3	FCz	77	68
	Subject 4	Cz	74	63
	Subject 5	FCz	78	63
Decision Tree	Subject 1	FCz	78	70
	Subject 2	FCz	77	70
	Subject 3	FCz	78	70
	Subject 4	FCz	76	68
	Subject 5	Cz	75	67

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
GMM	Subject 1	FCz	76	69
	Subject 2	Cz	72	66
	Subject 3	FCz	74	65
	Subject 4	FCz	66	64
	Subject 5	Cz	69	66

Table 3.8: Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the *EEG* protocol of P300 based BCIs.

### 3.6 Summary

In this chapter, three P300 based BCI protocols have been used to detect error related potentials. Our analysis show differences in ErrP waveforms across these three protocols. We have also tested the performances of ErrP classification using four different classifiers; SVM, LDA, GMM and Decision Tree classifiers. We observe that ErrP generated by *EEG* – which represents the highest level of user engagement among all three protocols – occurs much earlier than ErrP generated in other protocols. Our results also indicate that *EEG* protocol seems to have higher ErrP classification performance on average compared to other protocols. Figure 3.7 and Table 3.5 indicate that the difference in classification performance between *EEG* and the other protocols – *Observe* and *Control* – is statistically significant. These results suggest that ErrP waveforms and classification results are sensitive to user engagement in P300 based Brain Computer Interfaces.

## Chapter 4

# Analysis of ErrP in Motor Imagery based Brain Computer Interfaces

The movement of a limb or even the contraction of a single muscle can cause changes in brain activity around the cortex [37]. Even more so, mentally preparing for a movement or imagining such a movement has been shown to induce changes in sensorimotor rhythms (SMR). These rhythms, or oscillations, occur in the somatosensory and motor areas of the brain and are divided into five frequency bands. The delta band for frequencies less than 4 Hz, theta band for frequencies within the range of 4-7 Hz, alpha band for frequencies within the range of 8-12 Hz, beta band for frequencies within the range of 12-30 Hz, and the gamma band for frequencies above 30 Hz.

Sensorimotor rhythms can either increase (Event Related Synchronization, ERS) or decrease (Event Related Desynchronization, ERD) during the preparation or imagination of a movement. By capturing the ERD/ERS in the EEG and mapping them to a corresponding activity, left vs right hand motor imagery for example, it can be possible to train a classifier that is able to distinguish between the different movements. The alpha and beta bands are most frequently used in motor imagery classification.

This part of the thesis investigates the nature and detectability of error related potentials across three different motor imagery protocols described as follows.

1. MI Observe: The BCI automatically moves a ball on the screen while the user monitors its performance.

2. MI Control: The user presses a key on the keyboard to move a ball on the screen while noticing the response of the computer when the key is pressed.
3. MI EEG: The subject uses EEG to move a ball on the screen.

These three protocols can also be considered as a representation of three different forms of engagement for the user. One benefit that can be derived from this observation is if, after data analysis, user engagement enables us to explain the results obtained in each motor imagery based BCI protocol.

## 4.1 Processing Preexisting BCI Datasets

The first dataset used in testing our algorithms for error related potentials in the case of motor imagery based BCI has been collected by Millan and his colleagues [78]. Their study is based on the ability of detecting error related potentials in a motor imagery based BCI setting where the subject has no control of the system. This is similar to the *Observe* protocol we design in our experiments.

### 4.1.1 Experimental Description

In the experiment conducted by Millan et al., a cursor appears at the center of a screen with a target appearing either to the left or to the right of the cursor [1]. In a series of steps in either direction, the cursor attempts to reach the target during the course of a trial which lasts 2000 ms. The user has no control over these events but is instructed to observe whether the cursor moves in the desired direction, which is towards the target. For each step that the cursor makes, there is a probability  $P_{err}$  of the cursor moving in the opposite direction.

### 4.1.2 Results

Table 4.1 presents two sets of results; results obtained with ‘Our Code’ which is our implementation of the Gaussian Mixture Model Classifier described in Millan’s paper (see subsection 2.3.3). Following this are results presented in Millan’s paper

indicated by ‘Millan’s Code’. The similarity of our results with those reported in Millan’s paper provides a basis for using the GMM classifier we have designed in the our ErrP experiments.

Subject	Our Code		Millan’s Code	
	Correct(%)	Error(%)	Correct(%)	Error(%)
Subject 1	85	75	85	76
Subject 2	79	62	73	65
Subject 3	85	73	82	69
Subject 4	65	60	70	58
Subject 5	79	64	74	58
Subject 6	65	52	68	51

Table 4.1: ErrP classification results obtained by Millan et. al. [1] and the results we have obtained using our codes on the same data.

## 4.2 Designed Interface for MI ErrP

The entire system is written in C# and designed with the aid of Visual Studio. The system has two interfaces; a start menu and the actual experimental interface. The start menu, as seen in Figure 4.1, is used to acquire user information to keep a record. The start menu is also the same as seen in [79]. Following this, the train button is pressed and the main interface appears on the screen.



The screenshot shows a software window titled 'Information'. Inside, there are several input fields: 'Name' (a single text box), 'Birth Date' (three separate dropdown menus for day, month, and year), 'Gender' (a dropdown menu), 'Sampling Rate (Hz)' (a dropdown menu), and 'Electrodes' (a single text box). Below these fields, a 'Date' label is followed by the text '16 Agosto 2013 Cuma'. At the bottom of the window, there is a 'Ready.' label and two buttons labeled 'TRAIN' and 'TEST'.

Figure 4.1: The Start Menu of the Motor imagery related experiments.

The entire experiment is set up to include 3 runs, each of which contains the same number of trials. These trials are all identical in nature and can be split into three parts of chronological order.

1. A Resting Phase, where the subject can relax and prepare for the task at hand. This segment is brief and lasts 2 seconds. During this phase, the ball is blank and colored green.

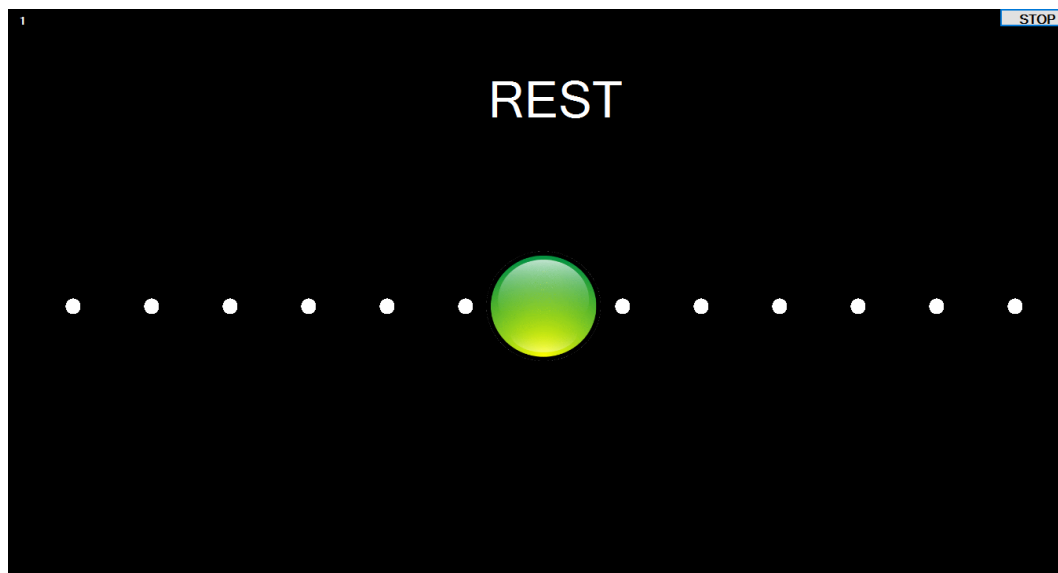


Figure 4.2: The Resting Phase of the Motor Imagery related experiments.

2. A Stimulus Phase, where the ball turns yellow and a white arrow appears on top of the ball pointing either to the left or to the right. The direction pointed by the arrow represents the direction of correct movement.

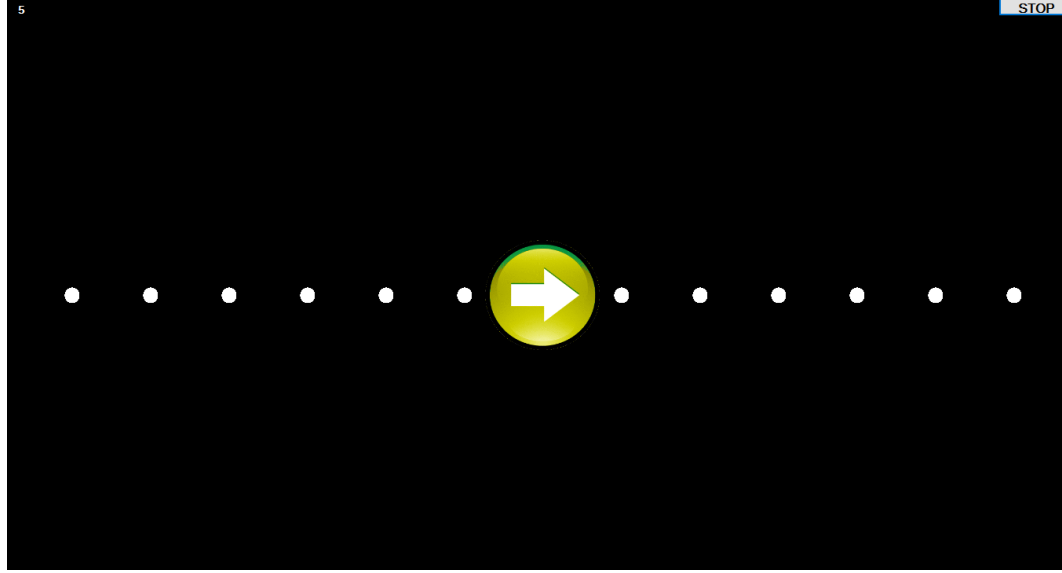


Figure 4.3: The Stimulus Phase of the Motor Imagery related experiments.

3. A Feedback Phase, where the the ball moves one step either to the right or to the left, and the subject is instructed to observe whether the ball moves in the correct direction or not. The ball moves in the correct direction with a certain probability  $P_{corr}$ . During this phase, the ball turns back to its original state of being blank with a green color.

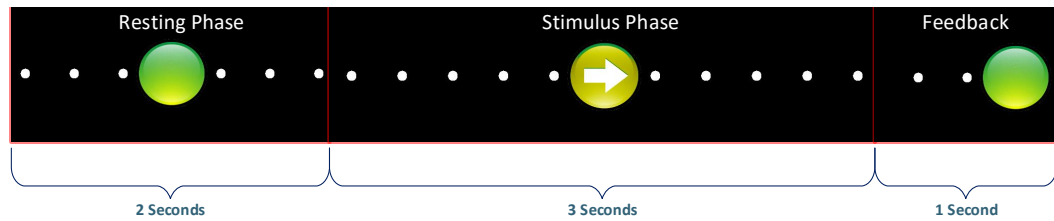


Figure 4.4: One trial in the motor imagery based BCI experiments.

### **4.2.1 Data Processing**

In these series of experiments, ErrPs are classified using EEG recorded from the FCz and Cz electrodes. To do this, data is collected in 8 different electrodes, namely, Fz, FCz, FC1, FC2, Cz, C1, C2, and CPz. Electrodes other than FCz and Cz are used for referencing. This is done by subtracting from an electrode the mean of data collected from the four adjacent electrodes, similar to the procedure in Chapter 3. After referencing, data from the FCz and Cz electrodes are downsampled to 64 Hz and then passed through a band pass filter with cutoff frequencies of 1-10 Hz. The sections of the data corresponding to the feedback phase were then obtained and categorized into two; error and correct samples. These are used as features in the classification process.

### **4.2.2 Classification**

The first two runs of each experiment were used as a training set and the last run as a testing set. The optimal classifier is then tested on the testing set and this process is performed 100 times for robust results. In this thesis, four different classifiers were investigated for ErrPs; SVM, LDA, Decision Tree, and a Gaussian Mixture Model classifier.

## **4.3 Motor Imagery Observe**

### **4.3.1 Experimental Description**

In this protocol, the stimulus phase lasts only one second. This is to ensure that the subject does not have enough time to use motor imagery. During the stimulus phase, the subject is instructed to notice the direction the arrow points to and observe if the ball moves in the right direction during the feedback phase.

### **4.3.2 Results**

Results in Table 4.2 were obtained based on the description provided in subsection 4.2.2.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	83.18	76.41	74.30	63.47
	Subject 2	82.72	75.30	81.40	73.55
	Subject 3	77.06	61.00	74.97	63.17
	Subject 4	77.92	71.00	74.10	66.05
	Subject 5	78.51	73.61	70.82	63.04
LDA	Subject 1	80.83	69.88	74.65	57.94
	Subject 2	81.77	71.15	80.42	69.35
	Subject 3	75.06	58.05	72.13	56.58
	Subject 4	79.20	67.60	75.65	61.15
	Subject 5	78.79	69.28	77.07	63.42
Decision Tree	Subject 1	79.25	69.94	77.20	67.74
	Subject 2	79.17	70.20	77.87	70.70
	Subject 3	74.53	67.11	75.34	66.11
	Subject 4	79.70	74.30	75.23	65.37
	Subject 5	76.41	68.76	74.79	65.52
GMM	Subject 1	72.44	67.17	71.60	67.11
	Subject 2	71.75	63.00	77.22	73.95
	Subject 3	69.27	58.41	70.00	66.70
	Subject 4	74.07	70.40	66.87	62.90
	Subject 5	65.92	62.95	60.30	54.76

Table 4.2: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the *Observe* protocol. For each electrode, the accuracy of classifying correct and error samples are presented.

## 4.4 Motor Imagery Control

### 4.4.1 Experimental Description

In this protocol, the duration of the stimulus phase is indefinite but subjects were advised to move the ball as quickly as possible. They do so by pressing the key 'A' to move the ball left and the key 'L' to move the ball right. The subject is instructed to press the key corresponding to the correct direction during the stimulus phase and to observe if the ball moves in the right direction during the feedback phase.

### 4.4.2 Results

Results in Table 4.3 were obtained based on the description provided in subsection 4.2.2.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	81.45	76.85	78.45	70.10
	Subject 2	80.09	74.21	76.17	64.36
	Subject 3	78.85	67.27	79.33	68.11
	Subject 4	82.95	75.94	75.85	63.50
	Subject 5	83.59	76.75	80.40	68.50
LDA	Subject 1	81.02	68.40	78.77	67.50
	Subject 2	78.17	63.94	73.60	60.94
	Subject 3	77.11	61.50	77.16	61.44
	Subject 4	82.26	71.44	74.59	56.33
	Subject 5	83.77	68.75	79.84	63.15
Decision Tree	Subject 1	77.47	70.40	76.85	69.80
	Subject 2	78.90	68.84	77.14	67.42
	Subject 3	75.85	67.33	77.07	67.55
	Subject 4	78.66	67.77	75.00	65.11
	Subject 5	84.11	74.50	81.34	67.62

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
GMM	Subject 1	72.30	66.75	70.15	65.20
	Subject 2	75.92	68.36	73.34	64.84
	Subject 3	71.26	65.83	71.33	65.16
	Subject 4	72.42	65.88	68.80	64.05
	Subject 5	79.97	69.06	74.45	69.00

Table 4.3: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the *Control* protocol. For each electrode, the accuracy of classifying correct and error samples are presented.

## 4.5 Motor Imagery EEG

### 4.5.1 Experimental Description

In this protocol, the stimulus phase lasts 3 seconds. During this time, the subject is instructed to use motor imagery to move the ball to either the left or to the right direction by imagining the respective arm movement. The users are told the movement direction is decided by the BCI system based on their EEG data, whereas in fact it is set randomly with a fixed probability of error. Following this, they are instructed to observe if the ball moves in the right direction.

### 4.5.2 Results

Results of the *EEG* protocol are provided in Table 4.4

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	78.57	68.88	69.16	59.11
	Subject 2	82.43	74.05	79.73	71.82
	Subject 3	76.06	64.17	74.04	60.41
	Subject 4	80.95	71.35	73.09	63.23
	Subject 5	76.97	71.76	70.23	65.00
LDA	Subject 1	76.40	62.38	76.50	61.44
	Subject 2	81.21	70.70	77.82	61.25
	Subject 3	76.37	60.88	76.97	60.41
	Subject 4	79.27	67.76	76.88	58.64
	Subject 5	76.84	69.61	75.71	62.95
Decision Tree	Subject 1	77.71	64.94	76.38	66.16
	Subject 2	81.46	74.58	77.56	70.58
	Subject 3	75.76	64.23	77.11	68.52
	Subject 4	77.16	67.52	76.41	64.23
	Subject 5	77.53	68.57	74.15	67.09
GMM	Subject 1	70.28	63.61	63.97	62.16
	Subject 2	79.65	70.52	75.41	68.35
	Subject 3	70.97	64.17	66.81	60.29
	Subject 4	73.74	70.05	65.79	62.88
	Subject 5	66.30	64.09	60.05	54.42

Table 4.4: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the *EEG* protocol. For each electrode, the accuracy of classifying correct and error samples are presented.

## 4.6 Discussion

### 4.6.1 Waveform analysis

Figure 4.5 shows the overall mean waveform for correct, error, and error minus correct samples. Blue lines represent average waveforms for correct samples, red lines represent average waveforms for error samples, and yellow waveforms represent the difference between the average error and average correct waveforms. The top row contains results obtained from the *Observe* protocol, the middle row is for the *Control* protocol, and the bottom row is for the *EEG* protocol. From left to right, the columns represent results for the FCz and Cz electrodes respectively.

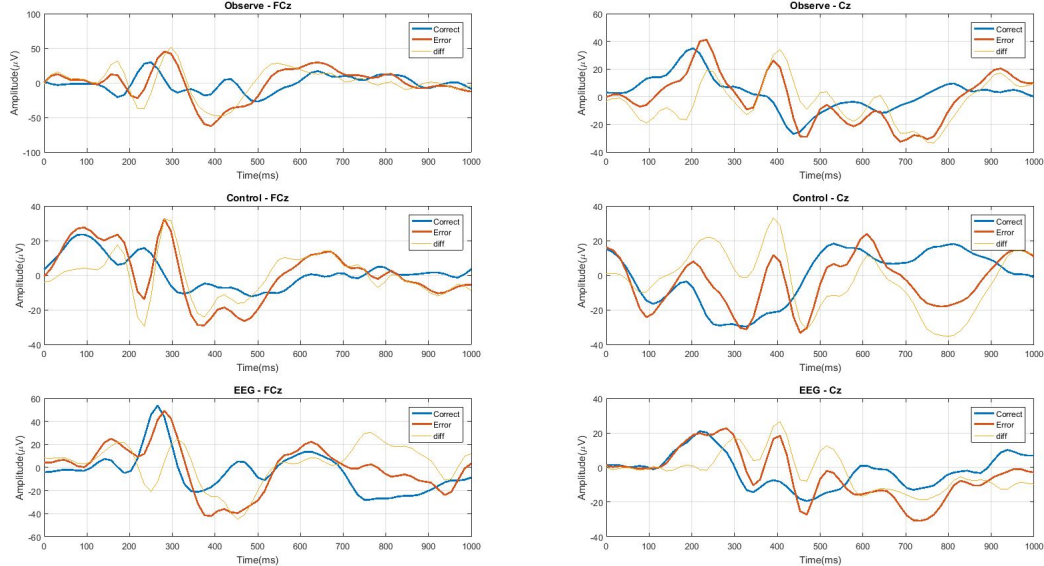


Figure 4.5: This figure presents overall average waveforms computed for correct (blue), error (red), and error-minus-correct (yellow) samples for two different electrodes and for all three protocols. From top to bottom, the rows are arranged in the following order: *Observe*, *Control*, *EEG*. The left column represents data from the FCz electrode and the right column represents data from the Cz electrode.

Figure 4.6 shows the mean error minus correct waveforms for all three protocols conducted. We have observed that the waveforms of all three protocols look similar and are consistent with the waveform obtained by Chavarriaga et. al. [1] (see Figure 2.6). The



similarities between the three waveforms were tested by first aligning the signals using a cross correlation procedure and then the correlation between each pairwise combination of error minus correct waveforms in Figure 4.6 is calculated. The correlation coefficient computed between the waveforms of *EEG* and *Observe* is  $\rho = 0.77$ ,  $\rho = 0.71$  for the waveforms of *EEG* and *Control*, and  $\rho = 0.88$  for the waveforms between *Observe* and *Control*.

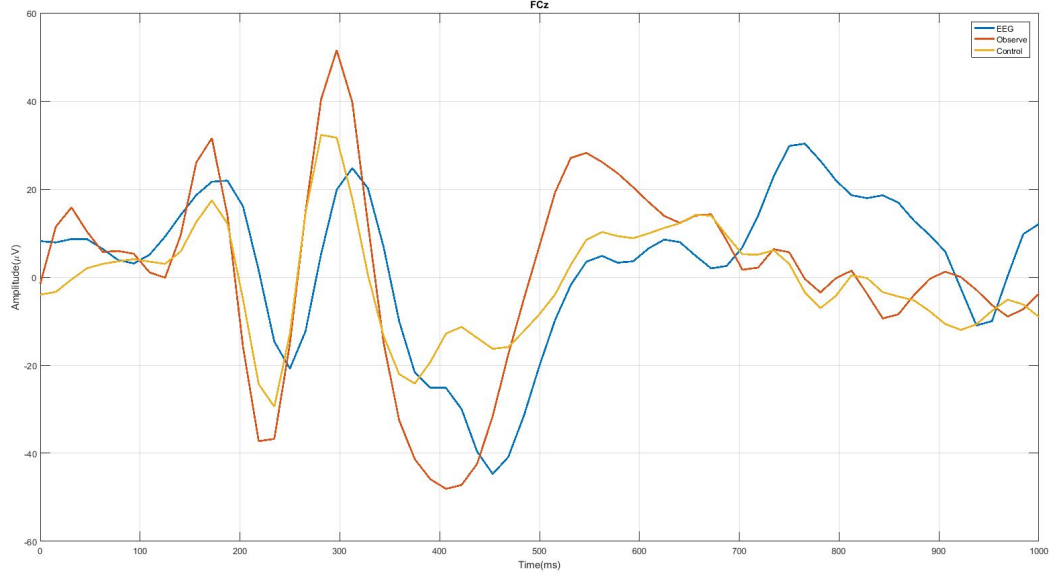


Figure 4.6: This figure shows the mean error minus correct waveforms for all three protocols recorded from the FCz electrode. Blue waveforms represent the *EEG* protocol, red represents *Observe* protocol and yellow represents the *Control* protocol.

In terms of the shape of the ErrPs, it can be seen that there is a small positive deflection at around 200 ms after feedback, followed by a negative deflection at 250 ms, another positive deflection at 300 ms followed by a prolonged negative deflection at 400 ms. This is consistent with results obtained by Ferrez et. al. [38]. Similar waveforms have also been reported by Spüler et. al. [80], Chavarriaga et. al. [78], and Kim et. al. [81]. Latency analysis shows that waveforms produced by *Control* and *Observe* have no phase difference. A statistically nonsignificant ( $p = 0.64$ ) phase difference of 15 ms exists between *Observe* and *EEG* protocols ( $p = 0.59$ ) as well as *Control* and *EEG* protocols ( $p = 0.28$ ).

The average maximum positive and negative deflections for all protocols based on Figure 4.6 are shown in Table 4.5. Pairwise statistical significance tests show that all peak amplitudes are statistically significant with  $p \ll 0.05$ .

Experiment Type	Max. Negative Deflection( $\mu V$ )	Max. Positive Deflection( $\mu V$ )
<i>Observe</i>	-48.06 ( $\pm 5.61$ )	51.52 ( $\pm 5.11$ )
<i>Control</i>	-29.44 ( $\pm 3.40$ )	32.28 ( $\pm 3.52$ )
<i>EEG</i>	-44.69 ( $\pm 6.92$ )	24.70 ( $\pm 6.67$ )

Table 4.5: Maximum positive and negative deflections computed from the waveforms provided in Figure 4.6.

### 4.6.2 Performance analysis

The offline classification results obtained in the previous section shows an above-chance level (50%) accuracy for all classifiers and subjects. To compare classifiers across experiments, we used the same method as used in subsection 3.5.2 and the results are presented in Figure 4.7. Classifier accuracies for all the optimal electrodes are contained in Tables 4.7, 4.8, and 4.9.

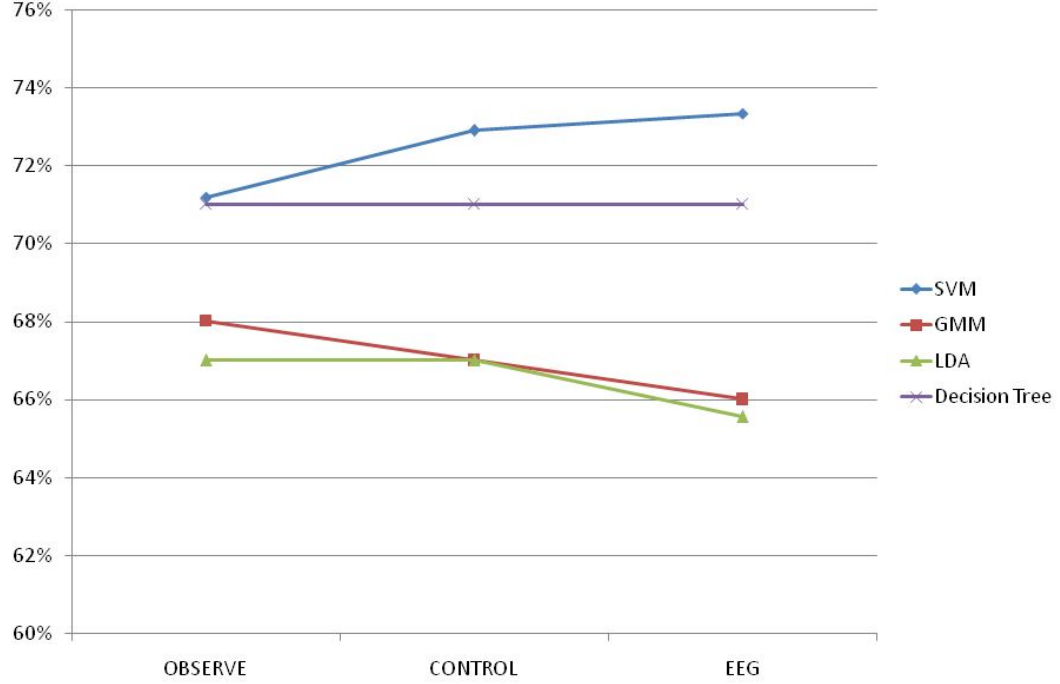


Figure 4.7: Average ErrP classification results over all subjects for all three protocols. Classifiers used include SVM, LDA, Decision Tree and GMM.

Figure 4.7 shows the mean classification results across subjects for each classifier. Observation indicates that SVM and Decision Tree classifiers perform better than LDA and GMM classifiers for all three protocols. Indeed, a t-test performed on the mean accuracies between all classifiers indicate that the difference between the performance of the SVM and Decision Tree classifier is not significant ( $p = 0.35$  for *EEG*,  $p = 0.57$  for *Observe*, and  $p = 0.41$  for *Control*). The difference between the performance of the LDA and GMM classifier is not significant either ( $p = 0.34$  for *EEG*,  $p = 0.50$  for *Observe*, and  $p = 0.58$  for *Control*), while all other combinations provide a significant difference with  $p < 0.01$ .

Across experiments, Table 3.5 provides the results of a MANOVA test on the average classifier performances of each pairwise combination of the three protocols.

Protocols \ Classifiers	Classifiers			
	SVM	LDA	Decision Tree	GMM
<i>Observe, Control</i>	<< 0.01	0.21	0.62	0.07
<i>Observe, EEG</i>	<< 0.01	<< 0.01	0.91	<< 0.01
<i>EEG, Control</i>	<< 0.01	0.01	0.77	0.32

Table 4.6: MANOVA test results for pairwise combinations of the three protocols in the motor imagery based experiments.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	83	73
	Subject 2	FCz	82	73
	Subject 3	Cz	74	63
	Subject 4	FCz	77	71
	Subject 5	FCz	78	71
LDA	Subject 1	FCz	80	69
	Subject 2	FCz	81	71
	Subject 3	FCz	75	58
	Subject 4	FCz	79	67
	Subject 5	FCz	78	69
Decision Tree	Subject 1	FCz	79	70
	Subject 2	FCz	78	72
	Subject 3	Cz	71	71
	Subject 4	Cz	73	70
	Subject 5	FCz	76	69

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
GMM	Subject 1	FCz	72	67
	Subject 2	Cz	77	73
	Subject 3	Cz	70	66
	Subject 4	FCz	74	70
	Subject 5	FCz	65	62

Table 4.7: Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the *Observe* protocol of motor imagery based BCIs.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	81	76
	Subject 2	FCz	80	74
	Subject 3	Cz	79	68
	Subject 4	FCz	82	75
	Subject 5	FCz	83	76
LDA	Subject 1	FCz	81	68
	Subject 2	FCz	78	63
	Subject 3	FCz	77	61
	Subject 4	FCz	82	71
	Subject 5	FCz	83	68
Decision Tree	Subject 1	Cz	80	72
	Subject 2	Cz	79	70
	Subject 3	FCz	75	67
	Subject 4	FCz	82	74
	Subject 5	FCz	84	74

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
GMM	Subject 1	FCz	72	66
	Subject 2	FCz	75	68
	Subject 3	FCz	71	65
	Subject 4	FCz	72	65
	Subject 5	FCz	79	69

Table 4.8: Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the *Control* protocol of motor imagery based BCIs.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	78	70
	Subject 2	FCz	71	76
	Subject 3	FCz	76	66
	Subject 4	FCz	80	75
	Subject 5	FCz	76	73
LDA	Subject 1	FCz	76	62
	Subject 2	FCz	81	70
	Subject 3	FCz	76	60
	Subject 4	FCz	79	67
	Subject 5	FCz	76	69
Decision Tree	Subject 1	Cz	76	66
	Subject 2	FCz	83	74
	Subject 3	Cz	77	65
	Subject 4	FCz	80	70
	Subject 5	Cz	79	71

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
GMM	Subject 1	FCz	70	63
	Subject 2	FCz	79	70
	Subject 3	FCz	70	64
	Subject 4	FCz	73	70
	Subject 5	FCz	66	64

Table 4.9: Optimal electrodes and ErrP classification performance of these electrodes for all subjects and classifiers in the *EEG* protocol of motor imagery based BCIs.

## 4.7 ErrP across P300 and motor imagery based BCI experiments

Analyzing error related potentials generated during P300 and motor imagery based BCI experiments simultaneously can be potentially interesting. Figure 4.8 shows the average error minus correct (difference) ErrP waveforms for both P300 and motor imagery based BCI experiments on the same plot for all three protocols. The blue line shows the difference waveform for motor imagery experiments scaled down by a factor of 4. The red line shows the difference waveform for P300 based BCI experiments without scaling. Since these averaged signals have been obtained from the same set of subjects, there seems to be an observable difference in ErrP waveforms generated between P300 and motor imagery based experiments. In the case of *Observe* and *Control*, waveforms generated from P300 and motor imagery based BCIs have a large negative deflection at around 400 ms. In the case of *EEG*, ErrP waveforms in the motor imagery case also has a negative peak at 400 ms whereas the ErrP waveform in the P300 case appears to be quite different. Another interesting observation is that waveforms generated in motor imagery based experiments have multiple positive and negative peaks in quick succession, while transition between peaks in P300 based waveforms seem to be much slower.

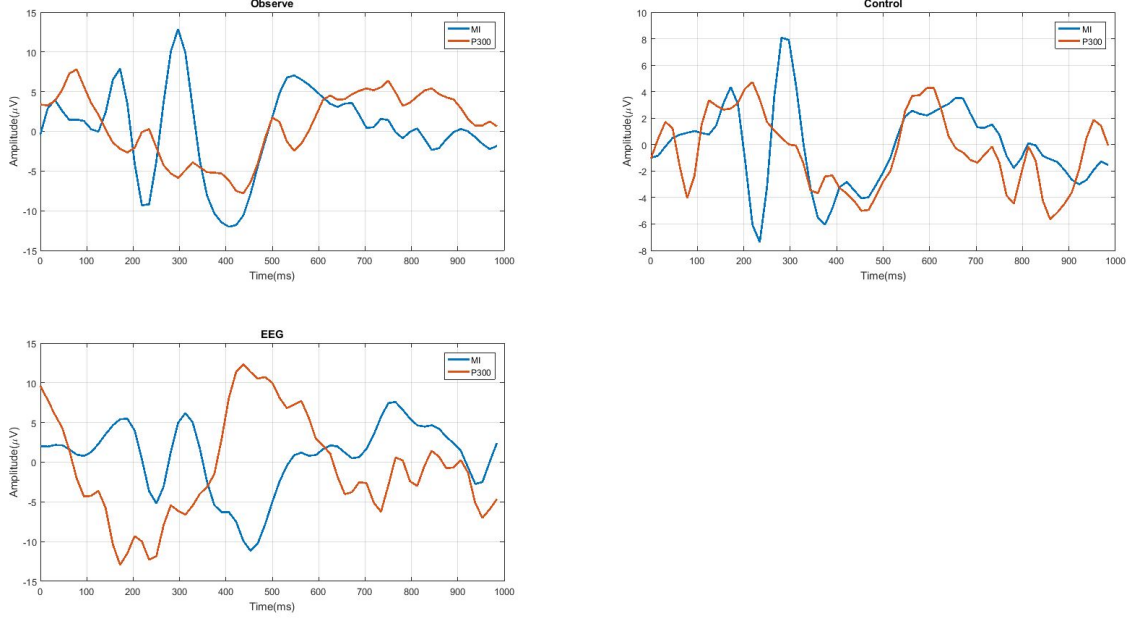


Figure 4.8: Average error minus correct ErrP waveforms for P300 and motor imagery based BCI experiments for all three protocols. Blue lines show the difference waveform for motor imagery experiments scaled down by a factor of 4. Red lines show the difference waveform for P300 based BCI experiments without scaling.

One more thing that could be potentially interesting is the consistency of ErrP signals with time for P300 and motor imagery based experiments. To analyze this, the average difference waveforms have been computed for all 3 runs separately and analyzed for all three protocols – *Observe*, *Control*, and *EEG* – in P300 and motor imagery based experiments. Figures 4.9 and 4.10 present these difference waveforms for motor imagery and P300 based experiments respectively. Blue lines represent difference waveforms obtained in the first runs, red lines correspond to waveforms obtained in the second runs, and yellow lines correspond to waveforms obtained in the third runs. An interesting result of this analysis is that for all motor imagery based BCI protocols, ErrP waveforms appear to be consistent as time passes. In contrast for P300 based experiments, there seem to be observable differences between the three waveforms obtained for any particular protocol. This consistency could perhaps explain why better classification results have been obtained in motor imagery based BCI protocols in comparison to the results obtained in P300 based experiments. These observations seem to indicate that



time does not have a dominant effect on stationarity of ErrP signals involved in our experiment.

Based on these results; one question that stands out is; to what extent does time have an effect on EEG nonstationarity? Additionally, what factors have a significant effect on nonstationarity in EEG and how can these factors be measured?

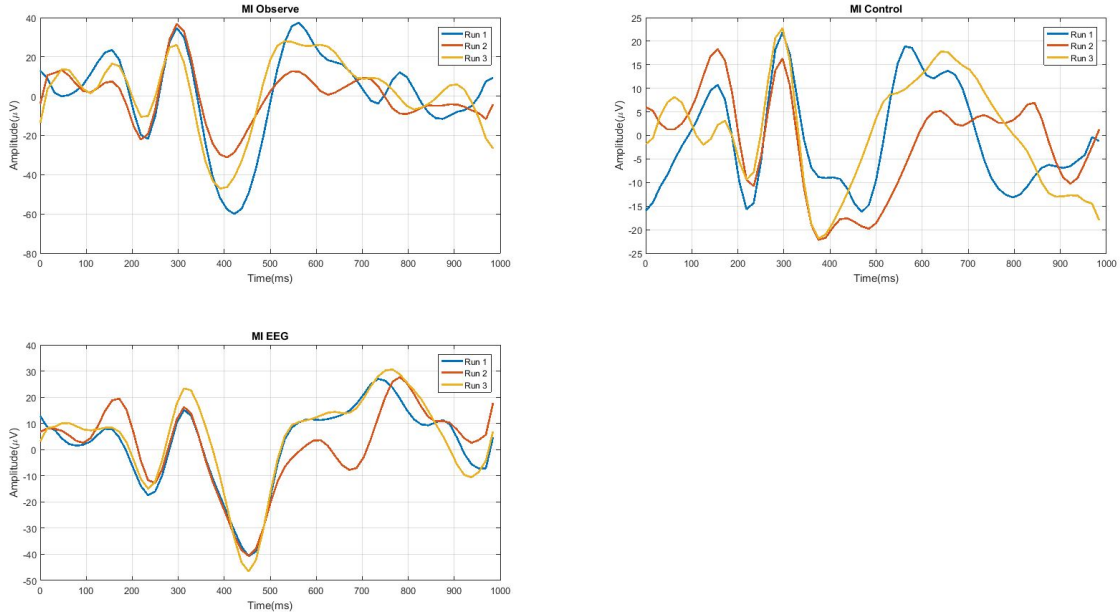


Figure 4.9: Difference waveforms of all three runs in motor imagery based experiments for all protocols. Blue lines represent difference waveforms obtained in the first runs, red lines correspond to waveforms obtained in the second runs, and yellow lines correspond to waveforms obtained in the third runs.

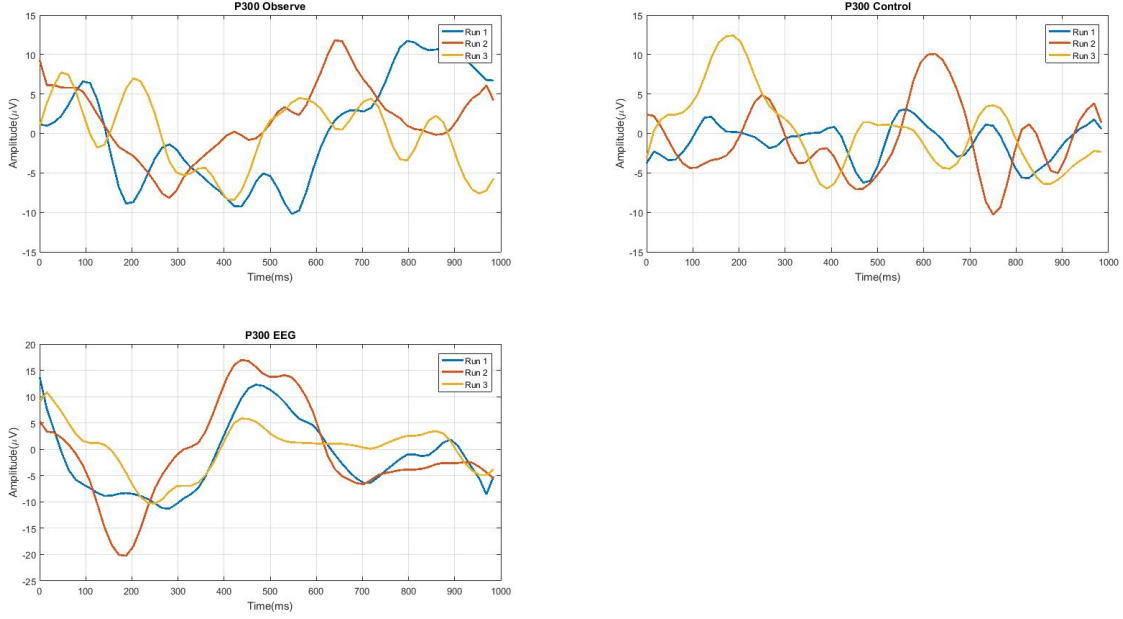


Figure 4.10: Difference waveforms of all three runs in P300 based experiments for all protocols. Blue lines represent difference waveforms obtained in the first runs, red lines correspond to waveforms obtained in the second runs, and yellow lines correspond to waveforms obtained in the third runs.

## 4.8 Summary

This chapter investigates error related potentials in three different motor imagery based BCI protocols; *Observe*, *Control*, and *EEG*. These protocols represent different contexts or user engagement levels. Analysis on all three protocols indicate that context has little or no effect on ErrP waveforms both in terms of latency and peak amplitudes. In fact, context also has little or no effect on ErrP classification performance as indicated by statistical significance tests performed on these results. These results are different from results obtained in Chapter 3 which leads us to believe that other factors need to be considered when determining how ErrP is affected in different contexts.

ErrP analysis across P300 and motor imagery based BCI experiments have also been conducted. ErrP waveforms generated in these experiments appear to differ significantly. The consistency of error related potentials have also been analyzed over

time for these two sets of experiments. It has been observed that while error related potentials generated for all motor imagery based BCI protocols remain consistent over time, the same cannot be said for the P300 based BCI experiments.

## Chapter 5

# Analysis and Design of a Joint Motor Imagery and ErrP-detection System

The initial problem we attempted to take on is finding a useful way of using ErrP to improve the performance of BCI systems. Two available options were motor imagery and P300 based BCI setups that have already been designed. Incorporating ErrP detection in both of these systems present different challenges.

For both systems, the first thing we wanted to focus on was the ability to detect the ErrP with a sufficient degree of accuracy – above the chance level of 50%. The second part of the problem is that given that an error related potential has been detected, how can this information be used to update a preexisting classifier? In the case of the motor imagery based BCI involving a binary classification problem, solving this problem is very easy. Assuming perfect ErrP detection, the presence of an error related potential means the user intended to move the cursor in one direction while the BCI misinterpreted this action and moved the cursor in the other direction. This information can be useful in adapting a classifier to minimize similar mistakes in the future.

In the case of a P300 based BCI however, we no longer have two choices but thirty-six. The consequence of this is that when an ErrP is been detected for any given trial, what we know with certainty is that the letter selected by the BCI is incorrect. If that

is the case, what is the true label? There are now thirty-five other options to choose from, instead of just one as in the case of the motor imagery based BCI. The multitude of these options is not the only challenging factor in this case. It is also the fact that how this information can be used to update the classifier is not very clear.

Based on the challenges with adaptation of P300 based BCIs, we decided to focus on ErrP based motor imagery BCI adaptation. This is a much simpler and straightforward. More importantly, working on it could potentially provide a better idea of what can be done when dealing with the P300 based BCI as well.

To do this, we develop a single motor imagery based BCI protocol that also accommodates the ability to detect error related potentials. The first step towards this goal is classifying the motor imagery and ErrP signals independently using the same interface. This leads to two preliminary experiments; a first set of experiments that only tested the detectability of ErrP in the system and a second set of experiments that only tested the performance of the motor imagery aspect of the system.

## 5.1 Detection of error related potentials

### 5.1.1 Design of Error Detection System

In this experiment, we aim to maximize the number of error related potentials that can be detected while minimizing the duration of an experiment as much as possible. While ErrP detection in Chapter 4 focused on errp analysis for various user engagement levels, this work investigates how errp detection is affected by the frequency of trials. To do this, three different paradigms have been designed.

1. **One Step Protocol:** In this setup, each trial starts with a green ball located in the middle of the screen. On either side of the ball, there is one target, a white dot, at the other end of the screen. When the *Start* button is clicked, the first trial begins with a 3 second rest period followed by a one second stimulus period. During this period, the ball turns yellow and an arrow appears pointing in the direction the ball is expected to move. This is followed by a feedback phase

during which the ball moves to an adjacent target. The BCI is programmed to move the ball in the wrong direction 30% of the time. This protocol has 12 trials per minute.



Figure 5.1: Interface of the one step protocol.

2. **Three Step Protocol:** In this setup, each trial starts with a green ball located in the middle of the screen. On either side of the ball, there are three targets, white dots, located equidistant from each other. When the *Start* button is clicked, the first trial begins with a 3 second rest period. This time, after the resting stage, the ball moves three times, with a one second resting period after each feedback phase. The BCI here is also programmed to move the ball in the wrong direction 30% of the time. This protocol generates about 16 trials per minute.

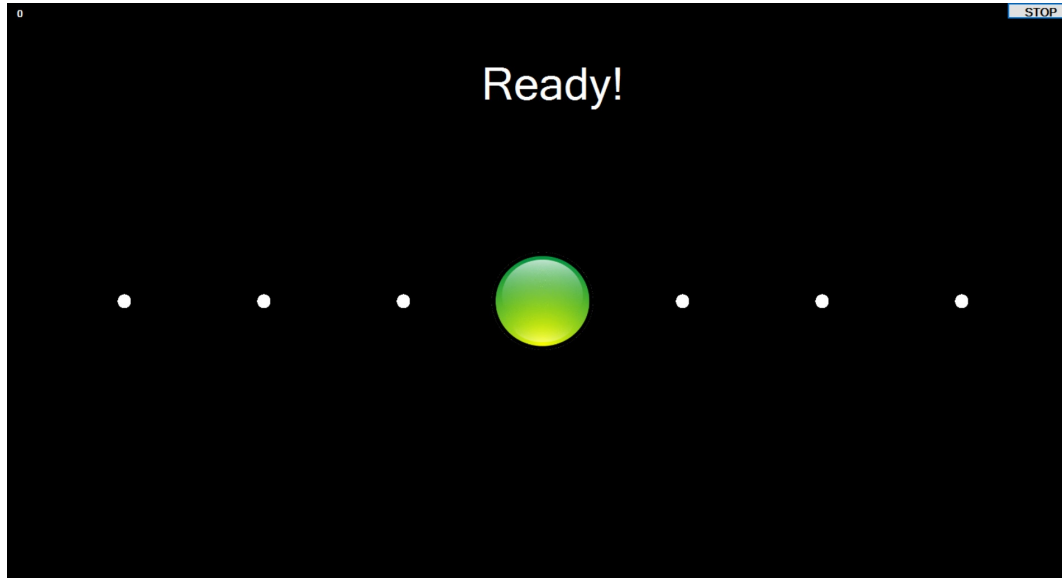


Figure 5.2: Interface of the three step protocol.

3. **Six Step Protocol:** In this setup, each trial starts with a green ball located in the middle of the screen. On either side of the ball, there are six targets, white dots, located equidistant from each other. When the *Start* button is clicked, the first trial begins with a 3 second rest period. This time, after the resting stage, the ball moves six times, with a one second resting period after each feedback phase. The BCI here is also programmed to move the ball in the wrong direction 30% of the time during the feedback phase. This protocol generates about 18 trials per minute.

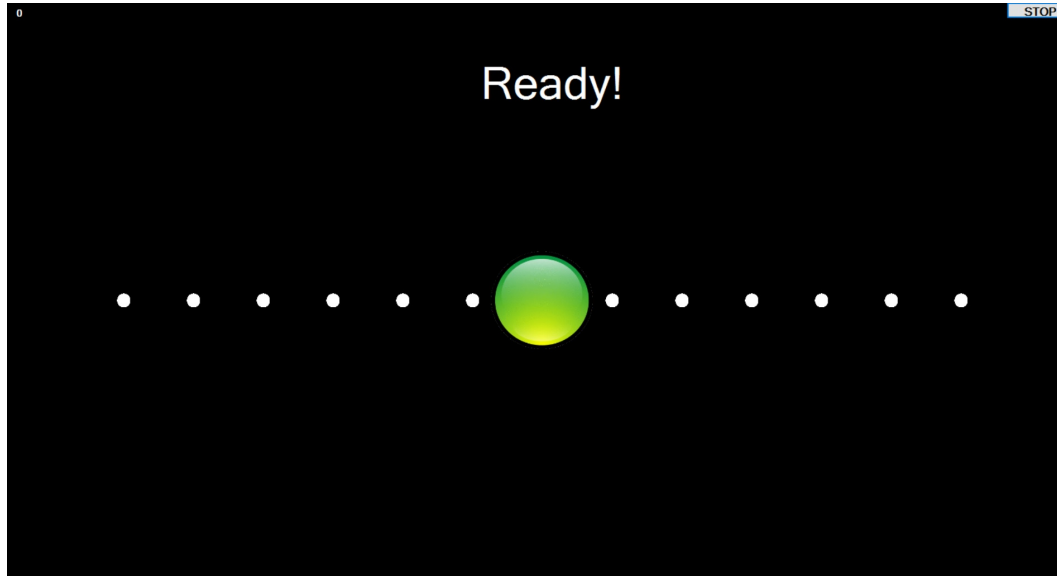


Figure 5.3: Interface of the six step protocol.

### 5.1.2 Results

Classification for all protocols were tested using four different classifiers; SVM, LDA, Decision Trees and Gaussian Mixture Model classifier. The results are shown in tables 5.1, 5.2, and 5.3. It can be observed that above chance accuracy has been achieved for all subjects in all classifiers.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	82	71	84	74
	Subject 2	82	64	82	71
LDA	Subject 1	82	64	84	68
	Subject 2	82	52	83	56
Decision Tree	Subject 1	83	66	86	74
	Subject 2	84	60	81	64
GMM	Subject 1	73	64	80	68
	Subject 2	78	67	74	69



Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)

Table 5.1: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the one step protocol.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	86	78	85	77
	Subject 2	86	74	80	76
LDA	Subject 1	82	71	83	71
	Subject 2	88	61	87	60
Decision Tree	Subject 1	85	76	82	75
	Subject 2	86	72	85	68
GMM	Subject 1	80	71	79	71
	Subject 2	73	65	73	60

Table 5.2: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the three step protocol.

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
SVM	Subject 1	88	66	83	59
	Subject 2	75	64	77	66
LDA	Subject 1	90	58	90	56
	Subject 2	76	58	74	60
Decision Tree	Subject 1	89	72	87	67
	Subject 2	76	70	76	66
GMM	Subject 1	82	70	80	71
	Subject 2	69	65	70	68

Classifier	Subject	FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)

Table 5.3: ErrP classification results for SVM, LDA, Decision Tree and GMM classifiers across all subjects for the six step protocol.

To compare the performances of different settings, we use a similar approach as used in subsection 3.5.2. The accuracies of the best performing electrodes in the training session are used as shown in tables 5.5, 5.6, and 5.7. The result of this analysis is shown in Figure 5.4. The blue line represents classification results for the One Step protocol, the red line is the results of the Three Step protocol, and the green line is the results of the Six Step protocol.

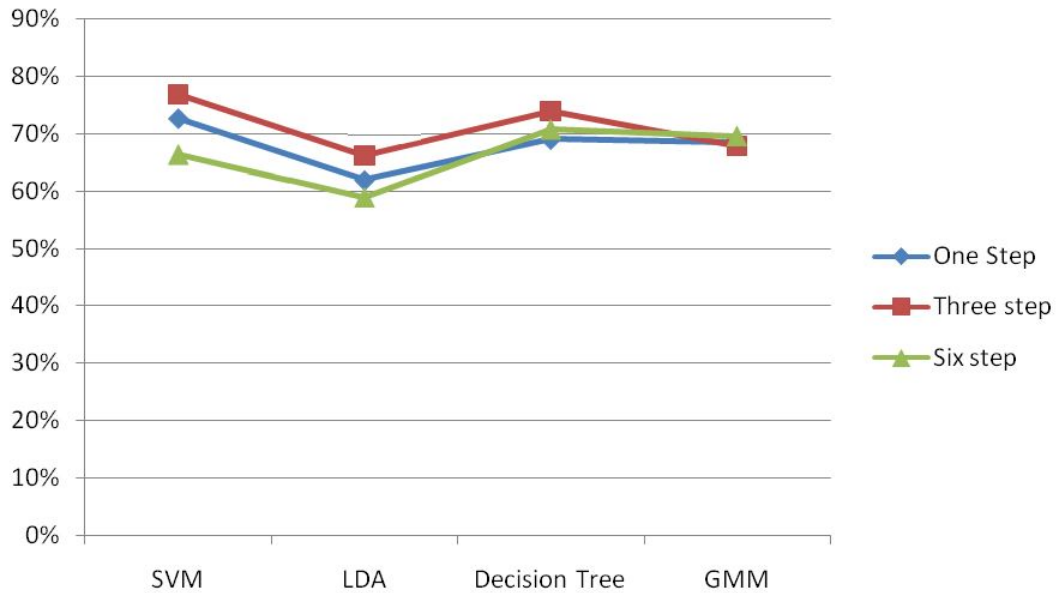


Figure 5.4: Average ErrP classification performance over all subjects for all three protocols. Classifiers used include SVM, LDA, Decision Tree and GMM classifiers.

Results presented in Figure 5.4 indicate that variations in performance is dependent on the classifier used. Indeed, MANOVA test performed for each classifier on the accuracies obtained for each protocol seems to confirm that observation. Table 5.4 confirms that SVM and LDA are volatile to changes in error frequency, Decision Tree is less volatile, while GMM seems to be robust to these changes.

Classifier	p-value	Decision
SVM	$p < < 0.01$	Significant
LDA	$p < < 0.01$	Significant
Decision Tree	$p = 0.39$	Not Significant
GMM	$p = 0.90$	Not Significant

Table 5.4: MANOVA test results for testing the significance of the difference between the accuracies obtained for the one step, three step, and six step protocols. This test is performed for four different classifiers; SVM, LDA, GMM, and Decision Tree.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	Cz	84	74
	Subject 2	Cz	82	71
LDA	Subject 1	Cz	84	68
	Subject 2	Cz	83	56
Decision Tree	Subject 1	Cz	86	74
	Subject 2	Cz	81	64
GMM	Subject 1	Cz	80	68
	Subject 2	Cz	74	69

Table 5.5: Optimal electrodes and their performances for all subjects and classifiers in the one step protocol.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	86	78
	Subject 2	Cz	80	76
LDA	Subject 1	Cz	83	71
	Subject 2	FCz	88	61

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
Decision Tree	Subject 1	FCz	85	76
	Subject 2	FCz	86	72
GMM	Subject 1	Cz	79	71
	Subject 2	FCz	73	65

Table 5.6: Optimal electrodes and their performances for all subjects and classifiers in the three step protocol.

Classifier	Subject	Preferred Electrode	Performance	
			Correct(%)	Error(%)
SVM	Subject 1	FCz	88	66
	Subject 2	Cz	77	66
LDA	Subject 1	FCz	90	58
	Subject 2	Cz	74	60
Decision Tree	Subject 1	FCz	89	72
	Subject 2	FCz	76	70
GMM	Subject 1	Cz	80	71
	Subject 2	Cz	70	68

Table 5.7: Optimal electrodes and their performances for all subjects and classifiers in the six step protocol.

## 5.2 Motor Imagery Classification

### 5.2.1 Interface Design and Data Processing

The interface for testing motor imagery is similar to the one used in Chapter 4. However note that EEG signals were not processed for MI classification in Chapter 4. The stimulus time was set to 6 seconds to accommodate motor imagery (see Figure 5.5) and during feedback, the ball always moves in the direction of the arrow to eliminate the effect of errors in the system.

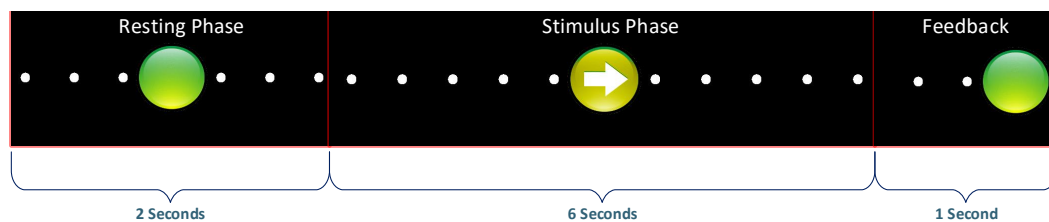


Figure 5.5: One trial in the preliminary motor imagery experiments.

The experiment was performed on 2 subjects and was split into 3 runs of 66 trials each, making a total of 198 trials. 5 minute breaks are provided between every run. The first two runs serve as the training set while the last run serves as the testing set. For motor imagery, data were collected from 7 electrodes; C3, Cz, C4, FC1, FC2, CP1, and CP2 electrodes at 2048 Hz. This was then downsampled to 128Hz. Each electrode was referenced with signals collected from the adjacent top, bottom, left, and right electrodes as shown in Figure 5.6. Spectral powers were calculated for all trials within three frequency bands; alpha,  $\alpha$ , 8-13Hz, beta  $\sigma$ , 14-18Hz, and beta,  $\beta$ , 19-24Hz. The mean power over these three bands for all three channels were used as features in an SVM classifier [79].

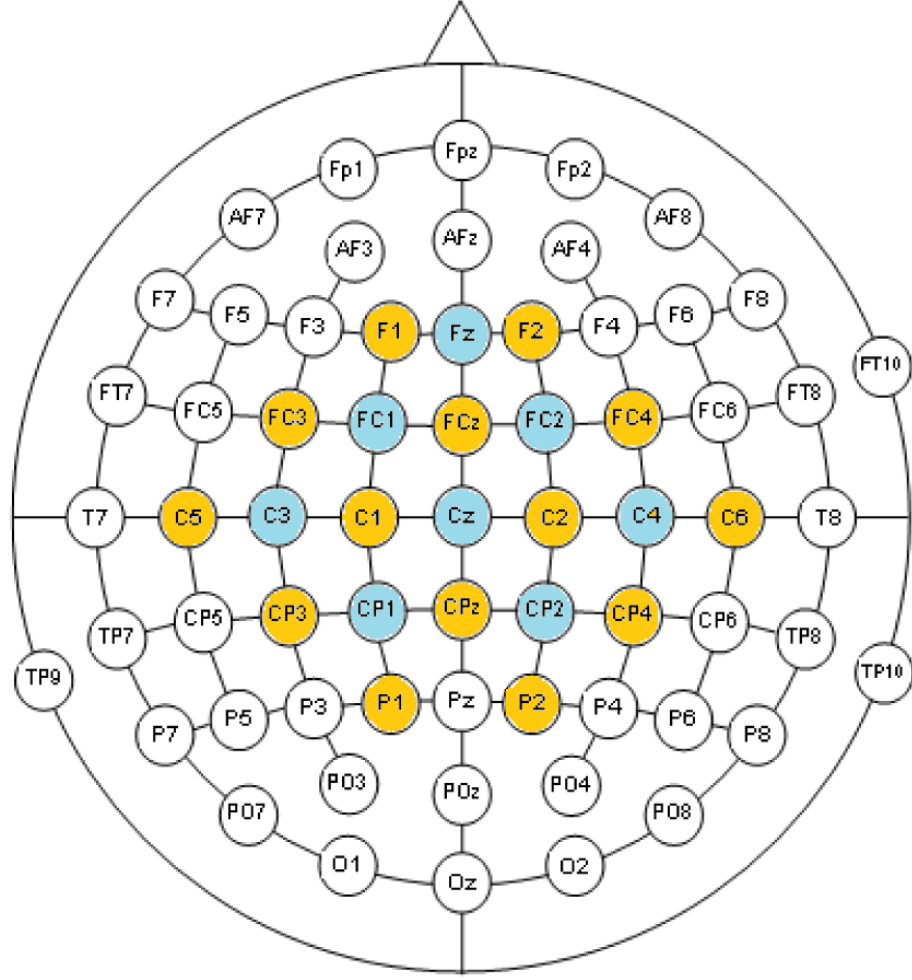


Figure 5.6: Electrodes used in the joint motor imagery and ErrP detection protocol. Blue colors represent electrodes used in classification and yellow represents electrodes used for referencing.

### 5.2.2 Results

Table 5.8 presents motor imagery classification results obtained for two subjects in the motor imagery experiment described in section 5.2. Both results are above 70% which is considered to be satisfactory in motor imagery based BCIs.

Subject No	Accuracy(%)
1	76
2	74

Table 5.8: Subject performance in motor imagery experiments.

## 5.3 Joint Motor Imagery and ErrP Detection

Results obtained in section 5.1 and section 5.2 have shown that it is possible to independently detect error related potentials and classify motor imagery on the interfaces we have designed. This provides a strong basis to attempt a joint classification of error related potentials and motor imagery in a single joint experiment. This preliminary work is presented in the following sections and has been tested on two subjects, neither of whom has participated in the previously described preliminary works.

### 5.3.1 Interface Design and Data Processing

The six step protocol as described in section 5.1 was chosen for the joint detection of ErrP and motor imagery. The stimulus time has also been set to 6 seconds as performed in the motor imagery case (see section 5.2).

The overall experiment contains 3 runs of 100 trials each with a 5 minute break in between each run. All run contains 100 9-second trials (see Figure 5.5). For each trial, the direction of the arrow is randomly chosen with equal probability for both sides. Motor imagery classification is performed offline so at the end of each trial, the ball moves in the wrong direction with a predetermined probability of 30%.

EEG was collected from 22 electrodes as shown in Figure 5.5. 7 electrodes – FC1, FC2, CP1, CP2, C3, C4 and Cz – were used for motor imagery classification and 2 electrodes – FCz and Cz – were used for ErrP classification. The rest were used for referencing these electrodes. Motor imagery classification was performed with the EEG collected in the 6-second stimulus phases and ErrP classification were performed with the EEG collected in the feedback phases. The same techniques used in chapter 4 and chapter 3 for motor imagery and ErrP classification.

## 5.4 Results

Table 5.9 shows the motor imagery classification accuracy obtained in the joint motor imagery and ErrP detection protocol. Table 5.10 shows the classification results obtained from FCz and Cz electrodes for all classifiers.

Subject No	Accuracy(%)
1	68
2	62

Table 5.9: Motor imagery performance in the joint motor imagery and ErrP detection experiments.

Classifier		FCz		Cz	
		Correct(%)	Error(%)	Correct(%)	Error(%)
LDA	Subject 1	78	61	78	58
	Subject 2	74	62	73	61
SVM	Subject 1	77	69	76	68
	Subject 2	76	67	75	66
Decision Tree	Subject 1	78	66	75	62
	Subject 2	74	69	76	69
GMM	Subject 1	68	65	68	63
	Subject 2	68	66	72	69

Table 5.10: ErrP classification results in the joint motor imagery and ErrP detection experiments.

Figure 5.7 also shows the error-minus-correct waveform averaged over all subjects over a period of 1 second. It starts with a positive peak right before the 200ms mark. This is followed by a negative deflection at around 250ms and then followed by another positive deflection at 300ms. The final negative deflection occurs at 400ms after feedback. This waveform is similar to the waveform obtained in the motor imagery *EEG*



protocol in shape but smaller in scale. A maximum peak of  $1.88\mu V$  and a maximum negative peak of  $-2.95\mu V$  can also be observed from the figure.

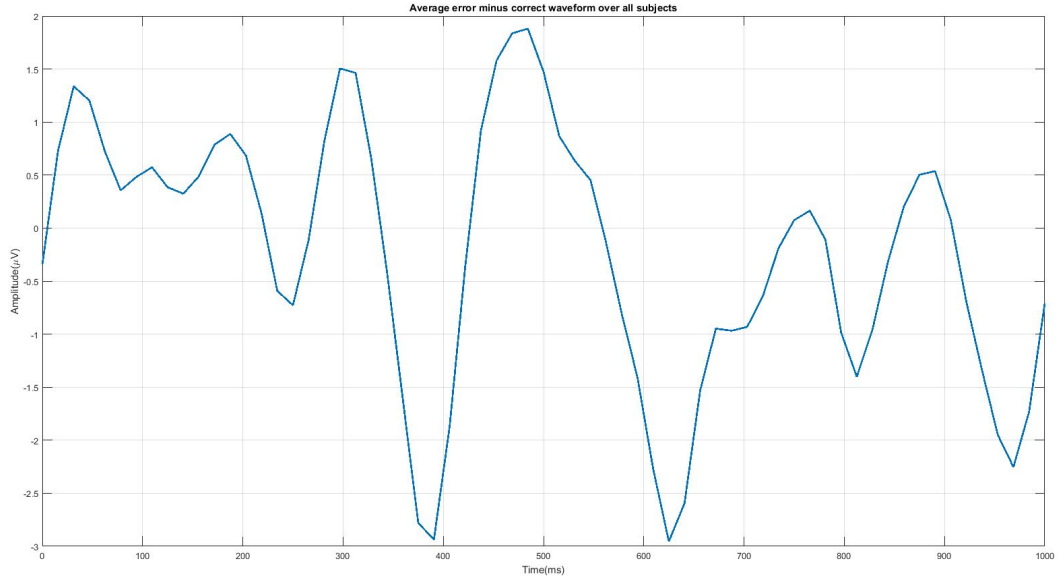


Figure 5.7: Error minus correct waveform averaged over all subjects for the FCz electrode.

Figure 5.8 shows the average classification performance over all subjects for four classifiers in the motor imagery *EEG* (blue line) and joint motor imagery and ErrP detection (red line) protocols. T-tests indicate significant differences ( $p < 0.01$ ) for LDA, SVM and Decision Tree classifiers and a nonsignificant difference for the GMM classifier ( $p = 0.29$ ).

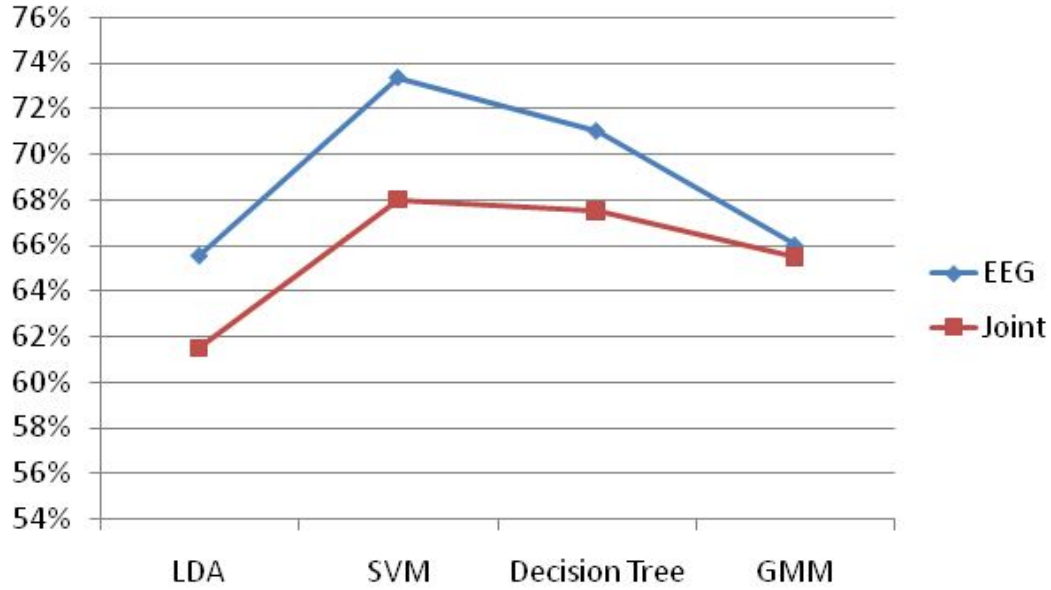


Figure 5.8: ErrP classification performance for LDA, SVM, Decision Tree, and GMM classifiers averaged over all subjects for motor imagery *EEG* (blue) and the joint motor imagery and ErrP detection (red) protocols.

Motor imagery results appear to be significantly lower when compared to results obtained in section 5.2. Similarly, ErrP classification results also appear to be lower when compared to the *MI-EEG* protocol. This could be due many reasons. In both cases, the joint motor imagery and ErrP detection system lasts much longer than either the motor imagery or *MI-EEG* protocols. Because of this reduced performance due to fatigue is very likely. Another factor is that for different subjects have been used across protocols. In fact, since only two subjects have participated in both the motor imagery and joint system, more subjects are needed to draw a firm conclusion.

## 5.5 Summary

In this chapter, three sets of preliminary studies have been conducted. The first is a study of how ErrP classification is affected by the frequency of trials in an experiment. The second is a study that tests motor imagery classification performance in our designed interface. In the third study, a BCI system that jointly classifies motor imagery

and error related potentials was tested.

Results of the first study shows that the effect of trial frequency on ErrP classification seems to be dependent on the classifier used. In the second study, motor imagery has been classified with accuracies comparable to those available in the literature. For the third study, results based on two subjects indicate that it can be possible to detect motor imagery and error related potentials simultaneously. It has also been observed that for both ErrP and motor imagery, average classification results seem to be decreased when compared to those obtained in the first and second preliminary studies.

## Chapter 6

# Conclusion and Future Work

### 6.1 Conclusion

In this thesis, we have analyzed error related potentials in P300 and motor imagery based BCI experiments. In particular, we have analyzed the effect of different types of user engagement on ErrP waveform and classification performance. In addition, we have made preliminary studies that analyze the effect of changes in frequency of trials on ErrP classification performance in motor imagery based protocols. We have also proposed and implemented a joint motor imagery and ErrP classification system.

Chapter 3 investigated error related potentials in the context of three P300 based BCI protocols. In this work, we confirm the presence of ErrP in three different P300 based BCI protocols and have shown that ErrP waveforms are consistent with findings reported in the literature. We have also observed that significant differences in ErrP waveforms and that differences in ErrP classification results existed across different protocols

We have also observed differences in terms of latencies and peak amplitudes across different protocols. Little difference exists between latencies and peaks of *Observe* and *Control* protocols ( $p = 0.49$ ) however, significantly higher peaks were observed ( $p < 0.05$ ) for *EEG* protocols. For *EEG* protocol which represents the highest user engagement level, ErrP seems to be generated much earlier. Additionally for all classifiers, the average ErrP performance is higher in *EEG* and statistically significant for

SVM and Decision Tree classifiers. We therefore conclude that ErrP waveforms are sensitive to the type of protocol and classification procedure used in a P300 based BCI. In this case, we believe user engagement has a positive effect on the performance of ErrP classification.

When comparing classifier performances across all three protocols, it was observed that two classifiers, SVM and Decision Tree, produced consistently better results for all experiments than the other classifiers, that LDA produces the lowest performance, and that GMM performing somewhere in between. This leads us to conclude that SVM and Decision Tree classifiers could be better suited for data of this form. This makes sense because SVM and Decision Tree classifiers have greater flexibility compared to LDA when working with data with an unknown probabilistic model. We believe this makes them more effective in classifying ErrPs since data for each subject can have different distributions. This could perhaps also explain the performance of LDA classifiers since they assume a normal distribution of data samples. Ideally, GMM classifiers should have performances comparable to SVM and Decision Tree classifiers since they have the benefit of modelling a wide range of distributions. We believe lower performance of these classifiers is a result of EEG nonstationarity.

Chapter 4 looked into error related potentials in three motor imagery based BCI protocols. We confirm their presence in all three protocols and we have shown that they are in accordance with findings reported in previous works. In contrast to the case of the P300 based BCI, we have discovered that different protocols contribute to little or no differences in the nature and classification performance of ErrPs.

Statistically significant differences were observed in terms of latencies and peak amplitudes but no evidence was found that these differences contribute to classification performance. Statistical tests performed on ErrP accuracies over all three protocols show that no significant differences exist for all four classifiers ( $\rho > 0.10$  for all pairwise combination of the three protocols). This leads us to the interesting conclusion that in a motor imagery based BCI, ErrP waveform can be slightly affected by the protocol used, but this impact is not sufficient enough to affect the classification accuracy for any of the four classifiers used in this work. We also hypothesize that the difficulty

of an experiment also plays an important role. All subjects reported P300 based BCI experiments were much more difficult than the motor imagery experiments. We therefore conclude that the effect of subject engagement on ErrP classification performance depends on experiment difficulty.

Comparison of classifier performances across all three protocols have lead to similar results as in the case of the P300 based BCI. Highest classification accuracies come from SVM and Decision Tree classifiers and lowest performance is produced by LDA. This provides more evidence in favor of the conclusion that higher performances are provided by classifiers that are more flexible to changes in the nature ErrP.

We have also analyzed how error related potentials change across P300 and motor imagery based BCI experiments. Results of this analysis indicate that ErrP waveforms differ significantly across these experiments. The consistency of ErrP waveforms also appears to be dependent on the type of experiment. For motor imagery based experiments, ErrP waveforms appear to be consistent across the three runs for all three protocols. In contrast, there appears to be high variability in ErrP waveforms generated through the three runs of all P300 based BCI protocols. These results seem to suggest that time alone has little or no short term effect on EEG stationarity and that other factors dependent on the type of BCI application affect EEG stationarity over time.

In chapter 5, a series of preliminary studies was conducted to test the feasibility of classifying motor imagery and error related potentials in a single experiment.

In the first study, we designed three sets of experiments to determine the effect of small changes in trial frequency on ErrP classification performance. These three experiments are; the one step protocol with a trial frequency of 12 trials/min, the three step protocol with a trial frequency of 16 trials/min, and the six step protocol with a trial frequency 18 trials/min. Our results indicate that the effect on ErrP classification performance depends on the classifier. We have observed SVM and LDA classifiers are more sensitive to changes in trial frequency and Decision Tree and GMM classifiers are more robust to changes in trial frequency.

In the second study, we designed a motor imagery classification experiment based on the interface of the six step protocol. Results were collected from two subjects over

200 trials and satisfactory classification accuracies over 70% were achieved for both subjects. The first two studies provide a strong foundation for conducting the third study in this chapter.

The third study implements a system where motor imagery and ErrP are classified in a single experiment. This experiment has 300 trials and was conducted on two subjects. Results of this experiment indicated that motor imagery and ErrPs can be classified simultaneously in a single experiment. It has also been observed that motor imagery performance is decreased in this protocol. We believe this is due to fatigue induced because of the longer trial periods.

We also conclude from the analysis of the results presented in figures 5.4 and 5.8 that GMM classifiers are more robust to changes in experiment parameters compared to other classifiers. In Figure 5.4, results are presented for three different trial frequencies and Figure 5.8, results are presented for two different lengths of stimulus phase in the motor imagery experiments.

## 6.2 Future Work

In this thesis, we have designed three different settings in P300 and motor imagery based BCIs and we have analyzed how these different settings affect error related potentials. One major result of this study is that in all cases, ErrP is classified with an above chance level. However, these results are obtained through offline analysis. Hence, one potential line of study following this work is implementing and analyzing these experiments in an online BCI system.

This work extracted one feature from the FCz and Cz electrodes in all ErrP classification experiments. This can be followed up by an investigation of other possible features and feature combinations that include electrodes outside the anterior cingulate cortex.

This work can also be improved by developing adaptation techniques that use error related potentials and analyzing how these techniques perform with respect to different contexts. These can be applied to both P300 and motor imagery based Brain Computer

Interfaces.

We have also conducted a preliminary study on the joint detection of motor imagery and error related potentials. Collecting samples of motor imagery and ErrP simultaneously came at a cost of increased experiment time of 33% and a decrease in motor imagery performance. More work can be done in finding ways of reducing this time and improving motor imagery performance. Another direction can also be an implementation and analysis of a system that detects P300 and error related potentials in a P300 speller.

If the effect of user engagement on ErrP waveform and classification performance depend on other factors, such as task difficulty or geometrical nature of the BCI application, a potential future work can be a comparison of results for P300 and motor imagery based BCI experiments with the same level of difficulty.

Results of motor imagery based BCI protocols indicate that ErrP waveforms are similar across protocols and are consistent within protocols. An interesting application of this knowledge can be the use of data trained on one protocol and tested on another protocol. The result of this analysis could also increase understanding of the extent to which these protocols are similar, or how they differ.



# Bibliography

- [1] R. Chavarriaga and J. d. R. Millán, “Learning from EEG error-related potentials in noninvasive brain-computer interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 381–388, 2010.
- [2] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain–computer interfaces for communication and control,” *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [3] J. Mackay, G. A. Mensah, S. Mendis, and K. Greenlund, *The atlas of heart disease and stroke*. World Health Organization, 2004.
- [4] H. Mitsumoto, *Amyotrophic lateral sclerosis: a guide for patients and families*. Demos Medical Publishing, 2009.
- [5] R. P. Van Peppen, G. Kwakkel, S. Wood-Dauphinee, H. J. Hendriks, P. J. Van der Wees, and J. Dekker, “The impact of physical therapy on functional outcomes after stroke: what’s the evidence?” *Clinical rehabilitation*, vol. 18, no. 8, pp. 833–862, 2004.
- [6] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, T. M. Vaughan *et al.*, “Brain-computer interface technology: a review of the first international meeting,” *IEEE transactions on rehabilitation engineering*, vol. 8, no. 2, pp. 164–173, 2000.
- [7] Z. A. Keirn and J. I. Aunon, “A new mode of communication between man and his surroundings,” *IEEE transactions on biomedical engineering*, vol. 37, no. 12, pp. 1209–1214, 1990.

- [8] W. Lang, D. Cheyne, P. Höllinger, W. Gerschlagel, and G. Lindinger, “Electric and magnetic fields of the brain accompanying internal simulation of movement,” *Cognitive brain research*, vol. 3, no. 2, pp. 125–129, 1996.
- [9] L. A. Farwell and E. Donchin, “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials,” *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.
- [10] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, “An eeg-based brain-computer interface for cursor control,” *Electroencephalography and clinical neurophysiology*, vol. 78, no. 3, pp. 252–259, 1991.
- [11] L. A. Farwell and E. Donchin, “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials,” *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.
- [12] E. Buch, C. Weber, L. G. Cohen, C. Braun, M. A. Dimyan, T. Ard, J. Mellinger, A. Caria, S. Soekadar, A. Fourkas *et al.*, “Think to move: a neuromagnetic brain-computer interface (bci) system for chronic stroke,” *Stroke*, vol. 39, no. 3, pp. 910–917, 2008.
- [13] K. K. Ang, C. Guan, K. S. G. Chua, B. T. Ang, C. W. K. Kuah, C. Wang, K. S. Phua, Z. Y. Chin, and H. Zhang, “A large clinical study on the ability of stroke patients to use an eeg-based motor imagery brain-computer interface,” *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 253–258, 2011.
- [14] T. M. Vaughan, D. J. McFarland, G. Schalk, W. A. Sarnacki, D. J. Krusienski, E. W. Sellers, and J. R. Wolpaw, “The wadsworth bci research and development program: at home with bci,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 229–233, 2006.
- [15] F. Nijboer, E. Sellers, J. Mellinger, M. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. Krusienski, T. Vaughan *et al.*, “A P300-based brain-computer in-

- terface for people with amyotrophic lateral sclerosis,” *Clinical neurophysiology*, vol. 119, no. 8, pp. 1909–1916, 2008.
- [16] L. R. Hochberg, M. D. Serruya, G. M. Friebs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, “Neuronal ensemble control of prosthetic devices by a human with tetraplegia,” *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.
- [17] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler, “An meg-based brain–computer interface (bci),” *Neuroimage*, vol. 36, no. 3, pp. 581–593, 2007.
- [18] G. Dornhege, *Toward brain-computer interfacing*. MIT press, 2007.
- [19] N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer, “Principles of a brain-computer interface (bci) based on real-time functional magnetic resonance imaging (fmri),” *IEEE transactions on biomedical engineering*, vol. 51, no. 6, pp. 966–970, 2004.
- [20] E. Niedermeyer and F. H. L. da Silva, *EEG Recording and Operation of the Apparatus*, 5th ed. Lippincott Williams and Wilkins, 2005.
- [21] J. J. Vidal, “Toward direct brain-computer communication,” *Annual review of Biophysics and Bioengineering*, vol. 2, no. 1, pp. 157–180, 1973.
- [22] T. Elbert, B. Rockstroh, W. Lutzenberger, and N. Birbaumer, “Biofeedback of slow cortical potentials. i,” *Electroencephalography and clinical neurophysiology*, vol. 48, no. 3, pp. 293–301, 1980.
- [23] G. Pfurtscheller and F. L. Da Silva, “Event-related eeg/meg synchronization and desynchronization: basic principles,” *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [24] T. W. Berger, J. K. Chapin, G. A. Gerhardt, D. J. McFarland, J. C. Principe, W. V. Soussou, D. M. Taylor, and P. A. Tresco, *Brain-Computer Interfaces: An*

*international assessment of research and development trends.* Springer Science & Business Media, 2008.

- [25] A. Amcalar, “Design, implementation and evaluation of a real-time P300-based brain-computer interface system,” Master thesis, Sabanci University, 2010.
- [26] Biosemi, 64 channels surgical medium/small (red/yellow) 10/20 layout. [Online]. Available: [http://www.biosemi.com/pics/cap\\_64\\_layout\\_medium.jpg](http://www.biosemi.com/pics/cap_64_layout_medium.jpg)
- [27] H. Berger, “ber das elektrenkephalogramm des menschen. zweite mitteilung ü,” *J Psychol Neurol*, vol. 40, pp. 160–179, 1930.
- [28] G. Pfurtscheller and F. L. Da Silva, “Event-related eeg/meg synchronization and desynchronization: basic principles,” *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [29] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, “An EEG-based brain-computer interface for cursor control,” *Electroencephalography and clinical neurophysiology*, vol. 78, no. 3, pp. 252–259, 1991.
- [30] J. R. Wolpaw and D. J. McFarland, “Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 51, pp. 17 849–17 854, 2004.
- [31] T. Solis-Escalante, G. Müller-Putz, C. Brunner, V. Kaiser, and G. Pfurtscheller, “Analysis of sensorimotor rhythms for the implementation of a brain switch for healthy subjects,” *Biomedical Signal Processing and Control*, vol. 5, no. 1, pp. 15–20, 2010.
- [32] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. L. Da Silva, “Mu rhythm (de) synchronization and eeg single-trial classification of different motor imagery tasks,” *Neuroimage*, vol. 31, no. 1, pp. 153–159, 2006.
- [33] F. Lopes da Silva *et al.*, “Event-related potentials: methodology and quantification,” 1998.

- [34] W. Walter, R. Cooper, V. J. Aldridge, W. C. McCallum, and A. L. Winter, “Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain,” *Nature*, vol. 203, pp. 380–384, July 1964.
- [35] S. Sutton, M. Braren, J. Zubin, and E. John, “Evoked-potential correlates of stimulus uncertainty,” *Science*, vol. 150, no. 3700, pp. 1187–1188, 1965.
- [36] E. Donchin and D. Smith, “The contingent negative variation and the late positive wave of the average evoked potential,” *Electroencephalography and clinical Neurophysiology*, vol. 29, no. 2, pp. 201–203, 1970.
- [37] B. Graimann, B. Z. Allison, and G. Pfurtscheller, *Brain-computer interfaces: Revolutionizing human-computer interaction*. Springer Science & Business Media, 2010.
- [38] P. W. Ferrez *et al.*, “Error-related EEG potentials generated during simulated brain–computer interaction,” *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 923–929, 2008.
- [39] J. DiGiovanna, B. Mahmoudi, J. Fortes, J. C. Principe, and J. C. Sanchez, “Coadaptive brain–machine interface via reinforcement learning,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 1, pp. 54–64, 2009.
- [40] I. Iturrate, L. Montesano, and J. Minguez, “Robot reinforcement learning using eeg-based reward signals,” in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4822–4829.
- [41] I. Iturrate, R. Chavarriaga, L. Montesano, J. Minguez, and J. d. R. Millán, “Latency correction of error-related potentials reduces BCI calibration time,” in *6th Brain-Computer Interface Conference 2014*, no. EPFL-CONF-200114, 2014.
- [42] Y. Motomura, A. Takeshita, Y. Egashira, T. Nishimura, Y.-k. Kim, and S. Watanuki, “Inter-individual relationships in empathic traits and feedback-related fronto-central brain activity: an event-related potential study,” *Journal of physiological anthropology*, vol. 34, no. 1, p. 1, 2015.

- [43] B. Dal Seno, M. Matteucci, and L. Mainardi, “Online detection of P300 and error potentials in a BCI speller,” *Computational intelligence and neuroscience*, vol. 2010, p. 11, 2010.
- [44] M. Spüler, M. Bensch, S. Kleih, W. Rosenstiel, M. Bogdan, and A. Kübler, “On-line use of error-related potentials in healthy users and people with severe motor impairment increases performance of a P300-bci,” *Clinical Neurophysiology*, vol. 123, no. 7, pp. 1328–1337, 2012.
- [45] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, “Toward enhanced P300 speller performance,” *Journal of neuroscience methods*, vol. 167, no. 1, pp. 15–21, 2008.
- [46] T. Zeyl, E. Yin, M. Keightley, and T. Chau, “Partially supervised P300 speller adaptation for eventual stimulus timing optimization: target confidence is superior to error-related potential score as an uncertain label,” *Journal of neural engineering*, vol. 13, no. 2, p. 026008, 2016.
- [47] —, “Adding real-time bayesian ranks to error-related potential scores improves error detection and auto-correction in a P300 speller,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 46–56, 2016.
- [48] C. S. Throckmorton, K. A. Colwell, D. B. Ryan, E. W. Sellers, and L. M. Collins, “Bayesian approach to dynamically controlling data collection in P300 spellers,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 3, pp. 508–517, 2013.
- [49] F. Jackson, B. D. Nelson, and G. Hajcak, “The uncertainty of errors: intolerance of uncertainty is associated with error-related brain activity,” *Biological psychology*, vol. 113, pp. 52–58, 2016.
- [50] A. Weinberg, R. Kotov, and G. H. Proudfit, “Neural indicators of error processing in generalized anxiety disorder, obsessive-compulsive disorder, and major depressive disorder.” *Journal of abnormal psychology*, vol. 124, no. 1, p. 172, 2015.

- [51] G. Hajcak, M. E. Franklin, E. B. Foa, and R. F. Simons, “Increased error-related brain activity in pediatric obsessive-compulsive disorder before and after treatment,” *American Journal of Psychiatry*, vol. 165, no. 1, pp. 116–123, 2008.
- [52] J. S. Moser, T. P. Moran, H. S. Schroder, M. B. Donnellan, and N. Yeung, “On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework,” 2013.
- [53] S. Theodoridis and K. Koutroumbas, *Pattern recognition*. Academic Press, 2009.
- [54] C. Bishop, “Bishop pattern recognition and machine learning,” 2001.
- [55] P. W. Ferrez and J. d. R. Millán, “You are wrong!—automatic detection of interaction errors from brain waves,” in *Proceedings of the 19th international joint conference on Artificial intelligence*, no. EPFL-CONF-83269, 2005.
- [56] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for eeg-based brain–computer interfaces,” *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007.
- [57] A. Y. Kaplan, A. A. Fingelkurts, A. A. Fingelkurts, S. V. Borisov, and B. S. Darkhovsky, “Nonstationary nature of the brain activity as revealed by eeg/meg: methodological, practical and conceptual challenges,” *Signal processing*, vol. 85, no. 11, pp. 2190–2212, 2005.
- [58] B. Blankertz, M. Kawanabe, R. Tomioka, F. Hohlefeld, K.-r. Müller, and V. V. Nikulin, “Invariant common spatial patterns: Alleviating nonstationarities in brain-computer interfacing,” in *Advances in neural information processing systems*, 2007, pp. 113–120.
- [59] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayouth, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, “A comparison of classification techniques for the P300 speller,” *Journal of neural engineering*, vol. 3, no. 4, p. 299, 2006.

- [60] R. C. Panicker, S. Puthusserypady, and Y. Sun, "Adaptation in P300 brain-computer interfaces: A two-classifier cotraining approach," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2927–2935, 2010.
- [61] P.-J. Kindermans, H. Verschore, D. Verstraeten, and B. Schrauwen, "A P300 BCI for the masses: Prior information enables instant unsupervised spelling," in *Advances in Neural Information Processing Systems*, 2012, pp. 710–718.
- [62] S. Lu, C. Guan, and H. Zhang, "Unsupervised brain computer interface based on intersubject information and online adaptation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 2, pp. 135–145, 2009.
- [63] T. Zeyl, E. Yin, M. Keightley, and T. Chau, "Adding real-time bayesian ranks to error-related potential scores improves error detection and auto-correction in a P300 speller," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 46–56, 2016.
- [64] B. D. Seno, "Toward an integrated P300 and ErrP based brain computer interface," Phd thesis, Politecnico di Milano, 2009.
- [65] P. Margaux, M. Emmanuel, D. Sébastien, B. Olivier, and M. Jérémie, "Objective and subjective evaluation of online error correction during P300-based spelling," *Advances in Human-Computer Interaction*, vol. 2012, p. 4, 2012.
- [66] Y. Li, H. Li, C. Guan, and Z. Chin, "A self-training semi-supervised support vector machine algorithm and its applications in brain computer interface," in *Acoustics, IEEE International Conference on Speech and Signal Processing, 2007. ICASSP 2007.*, vol. 1. IEEE, 2007, pp. I–385.
- [67] R. C. Panicker, "Adaptation and control state detection techniques for brain-computer interfaces," Ph.D. dissertation, 2011.
- [68] C. Vidaurre, A. Schlogl, R. Cabeza, R. Scherer, and G. Pfurtscheller, "A fully on-line adaptive BCI," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1214–1219, 2006.



- [69] I. Yilmaz, “Adaptation in P300 and motor imagery-based BCI systems,” Master thesis, Sabanci University, 2015.
- [70] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, “Co-adaptive calibration to improve bci efficiency,” *Journal of neural engineering*, vol. 8, no. 2, p. 025009, 2011.
- [71] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw, “Should the parameters of a bci translation algorithm be continually adapted?” *Journal of neuroscience methods*, vol. 199, no. 1, pp. 103–107, 2011.
- [72] P. W. Ferrez and J. d. R. Millán, “Simultaneous real-time detection of motor imagery and error-related potentials for improved bci accuracy,” in *Proceedings of the 4th international brain-computer interface workshop and training course*, no. LIDIAP-CONF-2008-019, 2008.
- [73] C. S. Carter, T. S. Braver, D. M. Barch, M. M. Botvinick, D. Noll, and J. D. Cohen, “Anterior cingulate cortex, error detection, and the online monitoring of performance,” *Science*, vol. 280, no. 5364, pp. 747–749, 1998.
- [74] M. Falkenstein, J. Hoormann, S. Christ, and J. Hohnsbein, “Erp components on reaction errors and their functional significance: a tutorial,” *Biological psychology*, vol. 51, no. 2, pp. 87–107, 2000.
- [75] C. B. Holroyd and M. G. Coles, “The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity,” *Psychological review*, vol. 109, no. 4, p. 679, 2002.
- [76] N. M. Schmidt, B. Blankertz, and M. S. Treder, “Online detection of error-related potentials boosts the performance of mental typewriters,” *BMC neuroscience*, vol. 13, no. 1, p. 1, 2012.
- [77] M. Spüler, W. Rosenstiel, and M. Bogdan, “Online adaptation of a c-vep brain-computer interface (BCI) based on error-related potentials and unsupervised learning,” *PloS one*, vol. 7, no. 12, p. e51077, 2012.

- [78] R. Chavarriaga and J. d. R. Millán, “Learning from EEG error-related potentials in noninvasive brain-computer interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 381–388, 2010.
- [79] E. Koyas, “Design and analysis of a brain-computer interface-based robotic rehabilitation system,” Master thesis, Sabanci University, 2013.
- [80] M. Spüler and C. Niethammer, “Error-related potentials during continuous feedback: using eeg to detect errors of different type and severity,” *Frontiers in human neuroscience*, vol. 9, p. 155, 2015.
- [81] S. K. Kim and E. A. Kirchner, “Handling few training data: Classifier transfer between different types of error-related potentials,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 320–332, 2016.